

Online hate speech: A study of YouTube comments on a criminal case of rape and homicide

Dimitra Filippou & Spyridoula Bella*

University of Patras, *National & Kapodistrian University of Athens

ΠΕΡΙΛΗΨΗ

Το παρόν άρθρο μελετά τη ρητορική μίσους και την κακοποιητική γλώσσα (abusive language), όπως πραγματώθηκαν σε σχόλια στο YouTube που σχετίζονται με μια ποινική υπόθεση που έχει προσελκύσει την προσοχή του κοινού και των μέσων ενημέρωσης στην Ελλάδα από το 2018. Υιοθετώντας μια κοινωνικοπραγματολογική προσέγγιση και συνδυάζοντας καταστασιακούς και κοινωνικούς παράγοντες με τα σχετικά νομικά πλαίσια, στοχεύουμε στον εντοπισμό στρατηγικών μέσω των οποίων πραγματοποιείται η ρητορική μίσους και η υβριστική γλώσσα μέσω συγκεκριμένων γλωσσικών πραγματώσεων. Για λόγους που σχετίζονται με τα υφιστάμενα νομικά πλαίσια, γίνεται διάκριση μεταξύ της ρητορικής μίσους και της κακοποιητικής γλώσσας για την αναφορά σε λεκτικές επιθέσεις κατά του θύματος και των δραστών, αντίστοιχα. Η ανάλυση των δεδομένων αναδεικνύει τον ρόλο που διαδραματίζουν το κοινωνικό πλαίσιο και η χρήση της γλώσσας στον καθορισμό του τι συνιστά ρητορική μίσους και πώς μπορεί να διακριθεί από άλλους τύπους προσβλητικής γλώσσας.

ΛΕΞΕΙΣ-ΚΛΕΙΔΙΑ: διαδικτυακή γλώσσα, έγκλημα, κακοποιητική γλώσσα, κοινωνιοπραγματολογία, ρητορική μίσους

1. Introduction

If one has a (not so thorough) look at the speeches of the rapporteurs of authoritarian and racist ideologies and/or leaders of totalitarian regimes, they will know that hate speech has always been a part of social life (see Filippou 2023, Guillén-Nieto 2023). In the past couple of decades, however, the phenomenon has attracted increasing legal and linguistic attention due to the significant expansion of online communication. According to Guillén-Nieto (2023: 33-34), the danger of online hate speech and the difficulty to combat it lies in factors like the anonymity of Internet users, the lasting duration of posts and comments, the possibility of multiple authorship of the same message from different accounts and the spreading of specific ideas across as many online spaces as possible for the sole purpose of inciting hatred and violence against individuals or groups, since users are free to roam across different platforms or social media.

Aiming at contributing to the understanding of the phenomenon and its linguistic manifestations, the present study focuses on hate speech emerging in YouTube comments made in relation to the high-profile criminal case of the double rape and murder of Eleni Topaloudi, a 21-year-old woman, in the Greek island of Rhodes in 2018. The night before her murder Topaloudi allegedly consented to have sexual intercourse with the two perpetrators. Despite following them to the residence of one of them, she refused intercourse. According to the pathologist's report, the perpetrators gang raped her, while beating her on the head with an iron object. Subsequently and while Topaloudi was unconscious, they carried her to the seashore and, after repeatedly beating her on the rocks, threw her into the sea where she met her death. Both perpetrators were convicted to a life sentence plus fifteen years for gang rape and intentional homicide.

To analyse the data of this study a sociopragmatic approach focusing on “the role of social conditions and variables in determining the use of language to mean and do things in the world” (Haugh et al. 2021: 4) is adopted. To this end, both the legal and the sociolinguistic context are exploited to allow the detection of specific strategies performing hate speech in the data, as well as typical linguistic manifestations of these strategies. For reasons explained in detail in Section 2.2 a distinction between hate speech and abusive language is made in reference to the data under examination. However, it

will be shown that this distinction, although consistent with legal approaches to hate speech, is disputable in light of the linguistic analysis to follow.

Therefore, the purpose of this study is twofold: (1) to explore the language of hate speech adding cross-linguistic evidence from Greek and (2) to problematize legal understandings of hate speech considering findings related to its linguistic manifestations. The paper is structured as follows. Section 2 presents the theoretical background of the study discussing both legal and linguistic perspectives, while Section 3 describes its method. The data is analysed in Section 4, while Section 5 discusses the findings and concludes the study.

2. Theoretical framework

This section presents the theoretical framework of the study, starting with legal and proceeding to linguistic approaches to the analysis of (online) hate speech.

2.1. Legal framework of hate speech

Hate speech as a legal concept has been of interest to scholars due to its possible legal consequences and its potential to trigger criminal acts (Tsesis 2002). The term was first introduced by Matsuda (1989), who attempted to describe the linguistic manifestations of racial hatred arguing that the right to freedom of opinion and expression should be both protected and limited by constitutional law. Brison (1998: 313) defines hate speech as “speech that vilifies individuals or groups on the basis of characteristics such as race, gender, ethnicity, religion and sexual orientation that (1) constitutes a personal insult, (2) creates a hostile or intimidating environment, or (3) is a form of group vilification” (see also Moran 1994). According to Waldron (2012: 6), hate speech is directly linked to attacks on human dignity by insulting, maliciously abusing, or defaming a section of the population (see also Ward 1997).

To date, there is no generally accepted definition of hate speech in the legal literature and the term itself has been considered “misleading” (Gelber 2017: 619), since “we know by experience that its use is neither limited to speech, nor to the expression of hatred” (Guillén-Nieto 2023: ix). More than any definition of the term, relevant treaties and conventions in the international, European, and Greek legal systems can be particularly enlightening regarding the legal aspects of hate speech.

Based on the Universal Declaration of Human Rights (UDHR) (1948), the International Convention on the Elimination of All Forms of Racial Discrimination (ICERD) places a responsibility on states to prohibit hate speech. The Council of Europe Decision 2008/913/JHA (28 November 2008) on combating forms and expressions of racism and xenophobia criminalises not only public incitement to violence or hatred, but also public condoning, denial, or gross trivialisation of crimes. Therefore, justifying and trivialising a crime against life, albeit not an incitement to violence, is considered hate speech. The most recent European initiative to curb hate speech is the Recommendation of the Committee of Ministers of the Council of Europe on 20 May 2022,¹ which calls on governments to develop comprehensive strategies to prevent and combat it, adopt an effective legal framework and implement appropriately calibrated and proportionate measures, ensuring that they are consistent with rights to freedom of expression and privacy.

¹ <https://www.coe.int/en/web/committee-on-combatting-hate-speech/-/the-council-of-europe-adopted-the-recommendation-on-combating-hate-speech>

Finally, Greek law specifies that “audiovisual media services shall not contain incitement to violence or hatred against a group of persons or a member of a group identified on the basis of the characteristics of race, colour, national or ethnic origin, descent, religion, disability, sexual orientation, identity or gender characteristics” (Article 8 based on Article 6 of Directive (EU) 2010/13).² These characteristics are collectively called “legally protected” and form the basis for all anti-racist and, consequently, hate speech legislation.

In general, existing legal frameworks provide some defining features for what constitutes hate speech. Specifically, it is public speech that incites violence against individuals or groups with “legally protected characteristics”³ and/ or violates fundamental rights of individuals or groups by infringing such characteristics.

2.2. Linguistic approaches to the analysis of (online) hate speech

The study of online hate speech from a linguistic perspective is a matter of increasing importance, since understanding the linguistic constitution of the phenomenon will (a) give new impetus and direction to the constantly evolving legal frameworks and (b) contribute the containment of the phenomenon and its consequences (Howard 2019, Paz María et al. 2020, Baider 2022, 2023).

Most studies on hate speech have employed corpus linguistics and computational tools to spot words and collocations through which it is performed and draw conclusions on machine and deep learning (see e.g. Biradar et al. 2021, Poletto et al. 2021, Davani et al. 2023). Rather few studies have sought to uncover linguistic features and strategies typical of hate speech. Knoblock’s (2022) edited volume on the *Grammar of Hate* examines morphology (Mattiello 2022), word formation (Beliaeva 2022), grammatical gender (Lind & Nübling 2022), the use of pronouns (Peterson 2022), the use of the imperative (Bianchi 2022) and specific syntactic structures (Geyer et al. 2022).

Guillén-Nieto’s (2023) book on hate speech focuses mainly on linguistic perspectives of the phenomenon. A major contribution of her involves the examination of hate speech in light of different theoretical approaches to language use (critical discourse analysis, genre, speech act theory, (im)politeness theory and cognitive pragmatics). Butler (1997: 19) explicitly introduces a crucial connection between hate speech and speech act theory (Austin 1975), which remains at the core of linguistic approaches to hate speech. She stresses that emphasis should be placed on the perlocutionary act that can result from the enactment of hate speech. Similarly, Guillén-Nieto (2023: 124) contends that:

“hate advocating speakers (a) say something by using words (locutionary act), (b) do something by using words – e.g. disseminate, advocate or incite hatred, hostility or violence against the members of a group identifiable by legally-protected characteristics (illocutionary act) and (c) try to engage the audience in unlawful or violent acts against the targets (perlocutionary act)”.

She proceeds to argue for the significance of the link between the illocutionary force and the perlocutionary effects and stresses that this link is “a core element in the legal approaches to hate speech” (ibid.). This view will be revisited in the analysis and discussion of the data under examination.

² Prohibition of Incitement to Violence or Hatred <http://www.opengov.gr/ggee/?p=6942>

³ Article 6 of Directive (EU) 2010/13.

Assimakopoulos et al.'s (2017) study constitutes, to our knowledge, the first linguistically oriented investigation of online hate speech. It does so from the perspective of critical discourse analysis, focusing on hate speech against immigration flows in the European Union. Among else, one major contribution of this study is the proposed distinction between *hard* and *soft* hate speech later referred to by Baider (2022, 2023) as *overt* and *covert* hate speech. While overt hate speech is realised through direct speech acts explicitly inciting violence, covert hate speech is latent and realised through indirect strategies such as metaphors (Musolff 2012), humour (Weaver & Bradley 2016), conspiracy theories (Baider 2022) and irony (Baider & Constantinou 2020). Consequently, while overt hate speech is criminalised under the provisions of international and European law, covert hate speech can pass undetected by non-linguists, despite sharing the same characteristics as overt (i.e., incitement to violence, spreading of hostility, violation/denial of legally protected human rights) and potentially leading to perlocutionary effects like those resulting from its overt counterpart. This makes online covert hate speech particularly hard to criminalise, since “the relationship between illocutionary and perlocutionary acts is not predetermined” (Bazerman 1994: 99) and the possible perlocutionary effects are by no means predictable. For this reason, the link between hate speech and the perlocutionary act is a key element in approaches and studies of the phenomenon (see Guillén-Nieto 2023).

Online hate speech has subsequently been studied by other researchers using different methodological approaches and analytical tools, such as Ruzaitė (2018: 93) who analyses hate speech in Lithuanian online comments and Technau (2018: 25) who studies ethnic and other group-based slur terms. Bick (2020) uses computational tools to investigate non-direct forms of hate speech, such as emoticons, while Becker's (2021) research on anti-semitism aims at contributing to finding ways to automatically detect hate speech in the online environment. Baider (2020, 2022) adopts a legal-linguistic approach, focusing on covert online hate speech and seeking to identify linguistic means employed by participants to counteract it.

In an in-depth study of the phenomenon and its relationship to identity construction and online conflict, Filippou (2023) seeks to identify the main strategies and linguistic manifestations of hate speech in YouTube comments related to three hate crimes that occurred in Greece in the past 8 years. Taking into consideration the existing legal frameworks and employing tools from sociopragmatics, she ends up with an inventory of four hate speech strategies:

- (1) justification and/or understatement of the crime,
- (2) insult to the memory of a deceased victim,
- (3) incitement to violence, and
- (4) dehumanisation.

These strategies are shown to be manifested by means of specific words, syntactic structures and pragmatic mechanisms that are partly context-dependent in terms of whether they are directed to the victim or the perpetrator(s), but are markedly similar irrespective of the crime they refer to.

The present study draws rather heavily on Filippou's (2023) findings, aiming to detect the strategies manifesting hate speech and their linguistic realisations in the data under examination. However, a certain idiosyncrasy of the Topaloudi case led us to opt for the term *abusive language* to refer to verbal attacks against the perpetrators of the crime and *hate speech* to attacks against the victim of the crime. The reason behind this choice is related to the fact that one of them was of Greek and the other of Albanian origin. According to the legal framework presented in Section 2, hate speech involves incitement to violence or hatred against a group of persons or a member of a group

identified on the basis of legally protected characteristics (along with justification, understatement and trivialisation of crimes against life). In the Topaloudi case the legal framework for legally protected characteristics could be triggered for the Albanian, but not the Greek perpetrator. Therefore, because the commentators treat the perpetrators as a single entity, the term abusive language is employed to describe public positionings against them that, although may not be, strictly speaking, hate speech, involve speech that can become harmful to society and trigger violence. The choice of the term *abusive* over related terms such as *offensive language* (see Culpeper & Haugh 2021: 200) is meant to emphasise the violent dimension of specific verbal attacks, since the data examined involve harsh, violent, or vulgar language and include explicit incitements to violence. It should be emphasised, however, that, despite its technical (i.e., legal) necessity, the validity of the distinction made here between hate speech and abusive language is disputable, since the data examined pose important questions as to the adequacy of the definitions of concepts associated with various forms of offensive language (cf. Janicki 2017: 157) and the linguistic analysis to follow highlights the possible need for reviewing legal conceptions of what constitutes hate speech.

3. Method

The data for the study were YouTube comments made in the comments section of three videos related to the rape and homicide of Eleni Topaloudi. The videos, informative in nature, were broadcasted on Greek television between 2018 and 2020 and were selected based on their popularity, i.e., the total number of views (>500,000).⁴ Their content focuses on issues connected to the crime (testimonies of the accused, course of the trial, statements of the victim's relatives etc.). The first and last 5 threads posted until April 2023 were selected from each video. This resulted in a dataset of 216 comments.

Data collection was carried out during the trial of the case. Specifically, the first-degree trial took place in May 2020, while the second-degree one ended in May 2022 with a unanimous guilty verdict for the defendants. The comments in Section 4 have been quoted unedited, i.e. the comments are reproduced verbatim and their translation into English has been done by the authors.

The data-driven analysis is qualitative. The MaxQDA programme for qualitative analysis was used to aid the detection of specific hate speech and abusive language strategies, as well as their most common linguistic manifestations. After determining what each comment attempted “to do” in the data, i.e. their illocutionary force, we ended up with two strategies performing genuine hate speech and another two which manifested abusive language. The coding of the linguistic manifestations of these strategies revealed specific patterns of language use. The strategies and their manifestations are analysed in the following section.

4. Data analysis

In our data, hate speech towards the victim of the crime was found to be manifested through justification and understatement of the crime and/or insulting the memory of a

⁴ YouTube links:

(1) https://www.youtube.com/watch?v=JhCWNN_c4nc

(2) <https://www.youtube.com/watch?v=GV5GDOMI12A>

(3) <https://www.youtube.com/watch?v=MsojoYwM2UU>

deceased person, while abusive language towards the perpetrators through incitement to violence and/or dehumanisation. Section 4.1 focuses on the former, while 4.2 on the latter.

4.1. Hate speech directed at the victim

4.1.1. Justification and understatement of crime

Based on the Council of Europe Decision 2008/913/JHA, justifying and trivialising a crime against life, albeit not an incitement to violence, is an insult to the victims. The justification of the double crime against the victim in the data functions as hate speech directed at her and is realised through specific linguistic mechanisms. T1's comment in example (1) is typical.⁵

(1) *Sorry alla den gnorizume agnostous stous dromous Kai tous kaloume gia sex spitia mas.. Kai oi dio plevres ftene gt katigoroume mono ekeinous, mipos epeidi einai alvanoï?*

[T1: Sorry but we don't meet strangers on the streets and invite them for sex in our homes. Both sides are to blame, why are we only blaming them? Is it because they're albanians?]

In (1), T1 initiates their contribution employing a structure of the “A... but B” type (*Sorry but...*), which functions as a hedging strategy to mitigate the second part of the utterance, i.e. the assertive speech act *den gnorizume agnostous stous dromous Kai tous kaloume gia sex spitia mas* (we don't meet strangers on the streets and invite them for sex in our homes) (cf. Holtgraves 1997).⁶ The *but* following *sorry* indicates that the speaker is not actually sorry for their statement. Moreover, the assertive speech act following *but* is expressed through the present tense denoting a high degree of certainty. As Bella (2005, 2007: 258) contends, the present tense due to the nature of its temporality (reference to the present) can express the non-temporal or the non-temporally bound. The use of the present tense by T1 in this utterance attributes to their statement a generic reference that is not temporally limited to “now” but presents their opinion as “part of the structure of the world” (Langacker 1987: 263), a norm that people follow as a rule. Hence, the opinion following *but* in the example expresses a deontic judgment,⁷ a rule or principle that being part of the structure of the world should be followed invariably. The use of the first-person plural, reminiscent of directives commonly addressed from parents to children (e.g., *πλένουμε τα χέρια μας πριν φάμε* ‘we wash our hands before eating’), reinforces this deontic reading, presenting T1's opinion as one of a wider group of people, i.e. as ‘intersubjective’ rather than ‘subjective’ (see Nuyts et al. 2010: 29). It follows that since the victim went against the norm, she is considered accountable and responsible for the crime. The justification of the crime and the victim blaming becomes explicit in the assertion *Kai oi dio plevres ftene gt katigoroume mono ekeinous* (‘both sides are to blame, why do we only blame them?’).

Similarly, in (2), the justification and understatement of the crime is explicit from the first utterance of T2:

⁵ The exact format and spelling with which they appeared on the platform was kept or all comments.

⁶ In this sense it is like constructions such as “I am not racist, but...”, which when used in specific contexts aim to obscure a racist positioning that follows (Geyer et al. 2022: 248).

⁷ We adopt here Nuyts et al. (2010) understanding of deontic modality in terms of the notion of (degrees) of moral or ethical acceptability or necessity, and not merely in terms of notions such as permission and obligation.

(2) και όλα αυτά γιατί η κοπέλα ήθελε να κάνει σχέσεις με υπανθρώπους αλλοδαπούς και κακά παιδιά

[T2: and all this because the girl wanted to have relationships with subhuman foreigners and bad boys.]

The use of the phrase *και όλα αυτά* ('and all this') to refer to the crime trivialises and understates it, while its justification is provided through the causal clause *γιατί η κοπέλα ήθελε να κάνει σχέσεις με υπανθρώπους αλλοδαπούς και κακά παιδιά* ('because the girl wanted to have relationships with subhuman foreigners and bad boys'). Given the subcategorization framework of the verb *θέλω* ('want'), the structure *η κοπέλα ήθελε να* (the girl wanted to), attributes to the victim the theta role of the experiencer⁸ (Chomsky 1981, Carnie 2006), i.e. the desire to be related to 'subhuman foreigners and bad boys', implying at the same time that the victim knew what she was getting into and thus was responsible for what she suffered. Assigning this thematic role to the victim creates a causal link with the outcome of the criminal act, presenting it, once more, as a "natural" consequence of the victim's rather than the perpetrators' behaviour.

4.1.2. Insult to the memory of a deceased person

According to the Criminal Code, "whoever insults the memory of the deceased by brutal or malicious insult or by slander shall be punished with imprisonment for up to six months" (Article 363). Based on this, insulting the memory of a deceased person is identified here as a manifestation of hate speech directed against the victim of the crime. Example (3) is typical:

(3) Καλό τανάκι ήταν κι αυτή

[T3: She was quite the slut too]

In (3), the word *τανάκι* (slut) derived from the word *πουτάνα* (prostitute/slut) to which the diminutive suffix *-άκι* (-aki) is added (*putána > putanáki > tanáki*) is a direct insult to the victim's memory. The shortened version of the word constitutes online slang, widely used on social media. This practice is rather common among social media users, when they seek to avoid explicit use of swearing or taboo language. However, it could also indicate that the speaker is aware that their positioning may have legal consequences, which they wish to avoid. This characterisation of the victim implies that she was responsible for the crimes against her. T3's comment is an instance of hate speech because: (a) it insults the victim on the basis of legally protected characteristics (gender and sexual life), (b) the victim is dead and Article 363 of the Criminal Code is triggered and (c) it constitutes a public positioning and not a private conversation to which there is no access and therefore no possibility for public hate spreading.

Example (4) is particularly interesting, since it combines the two main manifestations of hate speech examined in this section. Justification and understatement of the crime co-occurs with insult towards the memory of the victim:

(4) Όταν είσαι 19-20 χρόνων μέχρι ενός σημείου είναι λογικό να πειραματιστείς με κάποια πράγματα η να δοκιμάσεις κάτι για μια φορά η δεν ξέρω κ γω τι. Το δικαιολογεί κ η ηλικία. Αλλά επίσης η ηλικία είναι αρκετή για να αναλάβεις κ τις υποχρεώσεις σου και να μην προκαλείς την τύχη σου κάνοντας παρτουζα με 2 τύπους που μόνο που τους βλέπεις καταλαβαίνεις ότι δεν πάνε καλά. Ορισμένοι /ες εδώ μέσα γενικεύετε τις ευθύνες κ κατηγορείτε γενικά τους άντρες το οποίο θα το αφήσω ασχολίαστο γιατί δεν

⁸ An entity that undergoes an emotion, a state of being or a perception expressed by the verb.

θα κάτσω να λύσω εγώ τα κόμπλεξ σας απέναντι μας αλλά το μόνο που έχω να πω είναι ότι αυτοί ήταν καθάρματα και αυτή απρόσεκτη και με χωρίς ίχνος αυτό σεβασμού.

[T4: When you're 19-20 years old it's reasonable to experiment with some things or try something or I don't know what, your age justifies it. But you're also old enough to assume responsibility and not push your luck by having a threesome with two guys who you can tell just by looking at them that they are not in their right mind. Some people here are generalising the responsibilities and blaming men in general, which I will not comment upon because I am not going to sit here and solve your complexes towards us, but all I have to say is that they were scumbags, and she was careless and without a shred of self-respect].

The justification and understatement of the crime (see 4.1.1.) is manifested through the utterance *η ηλικία είναι αρκετή για να αναλάβεις κ τις υποχρεώσεις σου και να μην προκαλείς την τύχη σου κάνοντας παρτουζα με 2 τύπους* ('you're old enough to take care of your responsibilities and not push your luck by having a threesome with two guys'). Through this assertive speech act, T4 implicitly attributes responsibility for the crime to the victim because she did not 'take care of her responsibilities' and 'pushed her luck'. Furthermore, they underestimate a crime which has been judged as double rape by the judiciary, portraying it as a 'threesome'. The insult to the victim's memory comes through her characterisation as *απρόσεκτη και με χωρίς ίχνος αυτό σεβασμού* ('careless and without a shred of self-respect'). The fact that it is publicly directed towards the deceased victim of a hate crime with legally protected characteristics (gender) renders the characterisation a manifestation of hate speech.

4.2. Abusive language directed at the perpetrators

Following the rationale presented in 2.2 regarding the choice of the term abusive language, we have identified two main types of its manifestations in the data: (1) incitement to violence and (2) dehumanisation / inhumanisation. Dehumanization is the denial of a person's human status by depicting them in ways that devalues their humanity or individuality and, especially in the case of a prisoner, subjecting them to inhumane or degrading conditions (Bain et al. 2014, Over 2021, De Ruiter 2023). Inhumanisation involves the belief that certain individuals or groups are less human than others (Leyens et al. 2000). Both notions have been linked with extreme intergroup violence such as genocide (see e.g. Fyfe 2017). Incitement to violence and dehumanisation are exemplified and discussed in 4.2.1 and 4.2.2, respectively.

4.2.1. Incitement to violence

Public incitement to violence is an offence under Article 1 of Law 4285/2014. The content of the strategy is essentially the public dissemination of hatred, violent ideas and direct or indirect incitement to violence, regardless of the potential outcome of the speech act in question. The first part, the dissemination of hatred and incitement to violence, reflects the illocutionary act. The second part, namely the potential outcome that may result from the realisation of the speech acts, captures the perlocutionary act. In our data it involves encouraging violence and abuse towards the perpetrators. Example (5) is typical:

(5) *Θα φανε καλά από πίσω τα τομαρια!!!!*

[T5: They'll be fucked in the ass, these empty skins !!!!]

In example (5), T5 performs an assertive speech act expressing their belief that the perpetrators will be raped in prison. The dehumanising characterization *τομάρια* ('empty

skins’) licenses a deontic reading of the utterance. That is, it indicates that ‘being fucked in the ass’ is not simply what will happen to the perpetrators but what *should* happen to them *because* they are ‘empty skins’. This implicit causal link between dehumanisation and incitement to violence is a pattern identified in other cases of abusive language or hate speech (see e.g. Oberschall 2012 for such a link in the case of mass killings).

(6) *Οχι μονο να μη βγουν πρεπει να εκτελεστούν αλλά θα μου πεις δεν έχουμε νόμους εδώ.*

[T6: Not only should they not be released from prison, they must be executed, but we have no laws here.]

Example (6) constitutes a notable instance of incitement to violence, since the execution of the perpetrators is presented as a necessity and thus is explicitly encouraged by T6 (*πρεπει να εκτελεστούν* ‘they must be executed’). By means of the statement ‘but we have no laws here’ that follows, T6 most likely alludes to the death penalty, abolished in Greece since 1993 (Law 2172/1993). Both Amnesty International and the United Nations consider that the death penalty violates human rights, in particular the right to life and the right to live free from torture or cruel, inhuman, or degrading treatment or punishment (see Universal Declaration of Human Rights, 1948). Amnesty International declares that it “opposes the death penalty in all cases without exception - regardless of who is accused, the nature or circumstances of the crime, guilt or innocence or method of execution”. The death penalty has been judged to be one of the harshest acts of violence perpetrated by institutions and states against convicted criminals and has therefore been abolished within the borders of the European Union.⁹

Considering the legal framework in force in Greece and the European Union, the phrase *πρέπει να εκτελεστούν* (‘they must be executed’) functions as an incitement to violence against the perpetrators of the crime. The deontic modal *πρέπει* (‘must’) identifies the speech act as directive, presenting the violent death of the perpetrators as a necessity. Although the passive morphology of the verb *εκτελεστούν* (‘be executed’) conceals the agent of the execution, the statement *δεν έχουμε νόμους εδώ* (‘we have no laws here’) is indicative of the speaker’s view that imposing a violent death on the perpetrators should be a legitimate and legal course of action.

Similarly, (7) is a typical instance of public incitement to violence. The repetition of the noun *θάνατος* (‘death’) gives the comment the character of a slogan attributing the illocutionary force of an exhortation (‘kill the scumbag of society’) to both utterances:

(7) *Θάνατος στα αποβράσματα της κοινωνίας βιαστές και δολοφόνους!! Θάνατος!!*

[T7: Death to the scum of society rapists and murderers!! Death!!]

The choice of nominalisation, i.e. the use of the noun ‘death’ instead of the imperative *kill* conceals the agent of the act of killing ‘the scum of society’. However, it is inferable that the agents could be people with institutional roles (e.g. those who could bring back the death penalty), prison inmates or members of the public. It is worth mentioning that the abusive character of the comment is reinforced by the dehumanizing characterization ‘scum of society’, referring to rapists and murderers, in general, and the perpetrators of the crime under discussion in particular. This strategy is further exemplified in the following sub-section.

⁹ Since the 1990s, banning the death penalty has also become a condition for membership in the EU (see the Council of Europe website: www.coe.int).

4.2.2. Dehumanisation and Infrahumanisation

Compared to incitement to violence, dehumanisation can be considered a rather mild form of abusive language. To our knowledge, dehumanisation and infrahumanisation are not criminal acts and are not subject to any legal framework. Nevertheless, they can spread ideas that may become vastly harmful to society, such as incitement to discrimination, hostility, violence and crimes. As shown in 4.2.1 and is further exemplified in (8), dehumanisation can be used in conjunction with incitement to violence.

(8) *μόνο η ισόβια φυλακή τους αξίζει μέχρι την ημέρα που θα πεθάνουν αυτά τα απάνθρωπα τέρατα δεν έχουν το δικαίωμα ούτε το φως του ήλιου να δουν δεν πρέπει να βγουν ποτε απο τη φυλακή ούτε για ενα δευτερόλεπτο και για μένα σε όλους τους βιαστές και σε άλλα τέτοια τερατά που ζουν ανάμεσα μας τους αξίζει η θανατική ποινή*
 [T8: they deserve only life imprisonment until the day they die these inhumane monsters have no right to see the light of day, they should never get out of prison for a second and for me all rapists and other such monsters living among us deserve the death penalty]

In (8) dehumanisation is manifested through the reference to the perpetrators as *αυτά τα απάνθρωπα τέρατα* ('these inhumane monsters'). The characterisation *τέρατα* ('monsters') is repeated later in the comment (*τέτοια τέρατα [...] τους αξίζει [...] η θανατική ποινή* 'such monsters [...] deserve [...] the death penalty'). Thus, a causal connection is, once again, established between the portrayal of the perpetrators as non-human and the indirect incitement to violence encoded in 'they deserve the death penalty'.

The last example (9) was analysed in the subsection 4.1.1. as a manifestation of hate speech towards the victim of the crime. It is repeated here as an instance of infrahumanisation directed against only one of the perpetrators.

(9) *και ολα αυτα γιατι η κοπελα ηθελε να κανει σχεσεις με υπανθρωπους αλλοδαπους και κακα παιδια.*
 [T9: And all these because the girl wanted to have relations with subhuman foreigners and bad boys]

The speaker chooses to present the perpetrators as 'subhuman foreigners and bad boys'. The first characterisation refers to the Albanian, whereas the second at the Greek perpetrator. The infrahumanisation is manifested through the adjective *υπανθρωπους* ('subhuman'). It is interesting to note that, although both perpetrators committed the crime only one of them is infrahumanised ('subhuman foreigner'). Being directly related to the ethnic origin (legally protected characteristic) of the perpetrators, this distinction is one of the reasons why we argue our data lies on the boundary of abusive language and hate speech. This issue is taken up again in the discussion that follows.

5. Discussion and conclusions

This study examined YouTube comments made on a criminal case of rape and homicide to investigate online hate speech and abusive language. One of its main aims was to identify recurring strategies through which the two phenomena are manifested in the data, detect their typical linguistic realisations, and connect them to the context of the crime. Two main strategies of hate speech directed against the victim of the crime were identified in the data, namely justification and understatement of the crime and insult to the memory of the deceased victim. Abusive language, on the other hand, was found to involve incitement to violence and dehumanisation/infrahumanisation.

Tables 1 and 2 sum up the main linguistic manifestations for hate speech and abusive language, respectively, as they emerge from data analysis.

Table 1. Linguistic manifestations of hate speech (towards the victim)

<i>Linguistic Manifestations</i>	<i>Examples</i>
Insults	<i>She was quite the slut</i> (3)
Understatement of the crime	<i>not push your luck by having a threesome with two guys</i> (4)
Assertive speech acts with epistemic modality (increased certainty) and intention to justify the crime	<i>we don't meet strangers [...] and invite them [...] Both sides are to blame</i> (1)
Conventionalized syntactic structures with hedging function intended to justify or devalue the crime	<i>Sorry but we don't meet strangers on the streets and invite them for sex in our homes</i> (1)
Attribution to the victim of the thematic role of experiencer and agent instead of patient	<i>the girl wanted to build relationships with subhuman foreigners and bad boys</i> (2)

As shown in data analysis and partly repeated in Table 1 for the reader's convenience, the linguistic manifestations of hate speech can work in combination. For instance, a conventionalized syntactic structure with hedging function can lead to an assertive speech act with deontic connotations, through which the speaker justifies and devalues the crime, as in example 1. Furthermore, an insult to the victim can be combined with understating the crime (4). Therefore, the analysis of comments on this crime confirms Ruzaitė's (2018: 110) view that hate speech involves "creative language use", in the sense that speakers invent various linguistic means to keep it as covert as possible.

The linguistic manifestations of abusive language towards the perpetrators attested in our data and summarised in Table 2 can also co-occur.

Table 2. Manifestations of abusive language (towards the perpetrators)

<i>Linguistic Manifestations</i>	<i>Examples</i>
Phrases with vulgar content	<i>They'll be fucked in the ass</i> (5)
Dehumanizing (infrahumanizing) characterizations	<i>these inhumane monsters, such monsters (8), subhuman foreigners (9)</i>
Assertive speech acts with epistemic modality (increased certainty) and explicit reference to the use of violence	<i>They'll be fucked (5) such monsters deserve the death penalty (8)</i>
Directives speech acts with deontic modality (linked to incitement to violent acts)	<i>they must be executed (6) Death to the scum of society rapists and murderers (7)</i>

In example (5), for instance, a phrase with vulgar content is combined with an assertive speech act presenting the perpetrators' rape in prison as legitimate and just. Consequently, both hate speech and abusive language manifestations can emerge in combination and be mutually reinforcing.

A close comparison of the findings with those of Filippou (2023) reveals that both the strategies and the linguistic manifestations classified here as abusive language due to legal restrictions explained in Section 2.2 constitute genuine hate speech in the context of other crimes. That is, both incitement to violence and dehumanisation and their respective linguistic manifestations function as prototypical hate speech in different contexts. This brings us to the second aim of the study, i.e. the problematisation of the ways in which various types of offensive language are defined and treated by the law.

In relation to the distinction made between hate speech and abusive language, it was shown that, at least in the data under examination, the difference between the two phenomena is not so much a matter of strategies and linguistic manifestations, but one of context. Hate speech is identified as such, because it falls within a specific legal framework and, in our data, involves strategies and language aiming to verbally attack the victim of the crime. These manifestations fall within the category of hate speech because (a) the victim has legally protected characteristics, (b) the crimes against the victim have been adjudicated by the courts and, therefore, public questioning, belittling or denial of them is not justified and (c) the legislation protecting the memory of the deceased is triggered. While (a) reflects an understanding of hate speech frequently addressed in previous research (Assimakopoulos et al. 2017, Ruzaitė 2018, Baider 2020, 2022, Filippou 2023, Guillén-Nieto 2023), (b) and (c) have surfaced through the examination of the context of the crime in combination with legal frameworks and specific linguistic strategies attested in the data

What is worth noting is that the strategies classified as hate speech in our data do not express “hatred” in its usual sense. Instead, they portray the victim as contemptible and deserving of her violent fate. Hence, the danger in this case lies in the dissemination of opinions that devalue rape and murder, insult the memory of the victim and reinforce ideologies and discourses that may lead to new crimes in the future. This is in line with Butler’s view that hate speech can be understood “not only to act upon its listener (a perlocutionary scene), but to contribute to the social constitution of the one addressed (and, hence, to become part of a process of social interpellation)” (1997: 18). Such a process of interpellation entails a significant risk of generalisation to other victims and eventually to society’s perspectives on abominable crimes such as rape and murder. In other words, hate speech against victim(s) of crimes like the one examined here can perpetuate the view that abuse, rape and murder happen because of the victim’s behaviour and personality. In this sense, hate speech against the victim in this data can only be seen as covert.

Abusive language, on the other hand, was chosen here to refer to offensive and violence-triggering comments that target the perpetrators. As explained in Section 2.2, the case under examination is rather complex in that one of the perpetrators is not Greek and therefore the abusive comments attested in the data may border on hate speech due to his origin. Yet, the contextual parameter involving the Greek origin of the other perpetrator who does not have legally protected characteristics in the strict (legal) sense led us to distinguish between hate speech and abusive language.

Despite the technical definitions adopted here for hate speech and abusive language, the criminal case and the data of this study bring to the fore certain inadequacies of existing labels regarding language use.¹⁰ Although, by definition, hate speech can incite violence and trigger criminal acts, this does not –at first glance at least– appear to be the case in the comments characterised as hate speech in this study. That is, the comments against the victim in this case do not directly or explicitly incite violence. However, as already mentioned, they can promote beliefs, attitudes and ideologies according to which victims like Eleni Topaloudi are responsible for what happened to them and ultimately entrench such attitudes in public opinion

Comments targeted at the perpetrators and, for technical/legal reasons characterised as abusive language, on the other hand, involve multiple instances of explicit incitement to violence. The public character of such comments could potentially trigger multiple crimes such as torture, rape, harassment and physical assault. Therefore, although the

¹⁰ See also Janicki’s (2017) discussion about the problem of defining central and less central concepts associated with offensive language (hate speech, aggression, impoliteness etc.).

legal framework prevents us from labelling the language used against the perpetrators in these data as hate speech, this type of language, when used in public, can be equally harmful for society and promote violence and vigilantism. After all, the perpetrators of even the most abominable crimes have human rights that are protected by specific laws, even when they do not have legally protected characteristics as defined in hate speech legal frameworks. Therefore, public incitements to violence such as the ones attested in our data should be treated as (overt) hate speech and not merely abusive language.

The preceding discussion confirms Baider's observations about the insufficiency of existing legal frameworks to capture many instances of hateful and abusive speech that, as a result, "remain unaccountable before the law" (2023: 249). Acknowledging the enormous difficulties entailed in detecting hate speech in various contexts, Baider argues for "greater emphasis on counter-speech rather than censorship as the best way to deflect or halt hate speech" (id.). Specifically, she proposes the use of counter-narratives defined "as any form of expression that aims to influence those who sympathise with or take part in abusive speech" (Baider 2023: 250).

Although we could not agree more with Baider's views, we suggest that linguistic analyses like the one provided in this study, that investigate the ways in which meaning is generated through pairings of language use with specific situational and social contexts and legal factors should be taken into consideration when constructing legal frameworks on hate speech. Further research along these lines could shed light on the phenomenon, reveal its various aspects and, ultimately, contribute to its deflection. The red line between freedom of speech and penalising certain types of speech is, fortunately for linguists and unfortunately for legal scholars, a matter for the latter to figure out.

References

- Assimakopoulos, S., Baider, F. & Millar, S. 2017. *Online Hate Speech in The E.U. A Discourse-Analytic Perspective*. Cham, Switzerland: Springer.
- Austin, J. 1975. *How to Do Things with Words*. Oxford: Oxford University Press.
- Baider, F. 2020. Pragmatics lost? Overview, synthesis and proposition in defining online hate speech. *Pragmatics and Society* 11(2), 196-218.
- Baider, F. 2022. Covert hate speech, conspiracy theory and anti-semitism: Linguistic analysis versus legal judgement. *International Journal of Semiotics and the Law* 35, 2347-2371.
- Baider, F. 2023. Accountability issues, online covert hate speech, and the efficacy of counter-speech. *Politics and Governance* 11(2), 249-260.
- Baider, F. & Constantinou, M. 2020. Covert hate speech: A contrastive study of Greek and Greek Cypriot online discussions with an emphasis on irony. *Journal of Language Aggression and Conflict* 8(2), 262-287.
- Bain, P., Vaes, J. & Leyens, J. P. 2014. *Humanness and Dehumanization*. London: Routledge.
- Bazerman, C. 1994. Systems of genres and the enhancement of social intentions. In A. Freedman & P. Medway (eds) *Genre and New Rhetoric*. London: Taylor and Francis, 79-101.
- Becker, M. J. 2021. *Antisemitism in Reader Comments: Analogies for Reckoning with the Past*. London: Palgrave Macmillan.
- Beliaeva, N. 2022. Is play on words fair play or dirty play? On ill-meaning use of morphological blending. In N. Knoblock (ed.) *The Grammar of Hate. Morphosyntactic Features of Hateful, Aggressive, and Dehumanising Discourse*. Cambridge: Cambridge University Press, 177-196.
- Bella, S. 2005. Cognitive motivation and pragmatic functions of the Greek deictics. *Journal of Greek Linguistics* 6, 39-60.
- Bella: Μπέλλα, Σ. 2007. Ο ενεστώτας και η αντίληψη του παρόντος. [The present tense and the perception of the present]. Στο Τομέας Γλωσσολογίας (επιμ.) *Γλωσσικός Περίπλους*. Αθήνα: Καρδαμίτσα, 257-266.
- Bianchi, R. 2022. "Kill the invaders". Imperative verbs and their grammatical patients in Tarrant's The Great Replacement. In N. Knoblock (ed.) *The Grammar of Hate. Morphosyntactic*

- Features of Hateful, Aggressive, and Dehumanising Discourse*. Cambridge: Cambridge University Press, 222-240.
- Bick, E. 2020. Annotating emoticons and emojis in a German-Danish social media corpus for hate speech research. *RASK: International Journal of Language and Communication* 52, 1-20.
- Biradar, S., Saumya, S. & Chauhan, A. 2021. Hate or non-hate: Translation based hate speech identification in code-mixed hinglish data set. In *2021 IEEE International Conference on Big Data (Big Data)*, Orlando, FL, USA 2021, 2470-2475.
- Brison, S. 1998. The autonomy defense of free speech. *Ethics* 108(2), 312-339.
- Butler, J. 1997. *Excitable Speech: A Politics of the Performative*. London: Routledge.
- Carnie, A. 2006. *Syntax: A Generative Introduction*. Oxford: Blackwell.
- Chomsky, N. 1981. *Lectures on Government and Binding*. Berlin: Mouton De Gruyter.
- Culpeper, J. & Haugh, M. 2021. The metalinguistics of offence in (British) English: A corpus-based metapragmatic approach. *Journal of Language Aggression and Conflict* 9(2), 185-214.
- Davani, A., Atari, M., Kennedy, B. & Dehghani, M. 2023. Hate speech classifiers learn normative social stereotypes. *Transactions of the Association for Computational Linguistics* 11, 300-319.
- De Ruiter, A. 2023. To be or not to be human: Resolving the paradox of dehumanisation. *European Journal of Political Theory* 22(1): 73-95.
- Filippou: Φιλίππου, Δ. 2023. Γλωσσική αγένεια και κατασκευή ταυτοτήτων σε περιβάλλοντα συγκρούσεων και ρητορικής μίσους. [Linguistic Impoliteness and Identity Construction in Conflict Interactions and Hate Speech]. Διδακτορική διατριβή, Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών.
- Fyfe, S. 2017. Tracking hate speech acts as incitement to genocide in international criminal law. *Leiden Journal of International Law* 30(2), 523-548.
- Gelber, K. 2017. Hate speech definitions and empirical evidence. *Constitutional Commentary* 32, 619.
- Geyer, K., Bick E. & Kleene, A. 2022. "I am not a racist but...": A corpus-based analysis of xenophobic hate speech constructions in Danish and German media discourse. In N. Knoblock (ed.) *The Grammar of Hate. Morphosyntactic Features of Hateful, Aggressive, and Dehumanising Discourse*. Cambridge: Cambridge University Press, 241-261.
- Guillén-Nieto, V. 2023. *Hate Speech: Linguistic Perspectives*. Berlin: De Gruyter Mouton.
- Haugh, M., Kádár, D. & Terkourafi, M. 2021. Introduction: Directions in Sociopragmatics. In M. Haugh, D. Kádár & M. Terkourafi (eds) *The Cambridge Handbook of Sociopragmatics*. Cambridge: Cambridge University Press, 1-12.
- Holtgraves, T. 1997. Yes, but... positive politeness in conversation arguments. *Journal of Language and Social Psychology* 16(2), 222-239.
- Howard, J. W. 2019. Free speech and hate speech. *Annual Review of Political Science* 22, 93-109.
- Janicki, K. 2017. What is conflict? What is aggression? Are these challenging questions? *Journal of Language Aggression and Conflict* 5(1), 156-166.
- Knoblock, N. (ed.). 2022. *The Grammar of Hate: Morphosyntactic Features of Hateful, Aggressive, and Dehumanizing Discourse*. Cambridge: Cambridge University Press.
- Langacker, R. W. 1987. *Foundations of Cognitive Grammar. Vol. I: Theoretical Perspectives*. Stanford: Stanford University Press.
- Leyens, J. P., Paladino P., Rodriguez-Torres, R., Vaes, J., Demoulin, S., Rodriguez-Perez, A. & Gaunt, R. 2000. The emotional side of prejudice: The attribution of secondary emotions to ingroups and outgroups. *Personality and Social Psychology Review* 4(2), 186-197.
- Lind, M. & Nübling, D. 2022. The neutering neuter. The discursive use of German grammatical gender in dehumanisation. In N. Knoblock (ed.) *The Grammar of Hate. Morphosyntactic Features of Hateful, Aggressive, and Dehumanising Discourse*. Cambridge: Cambridge University Press, 118-139.
- Matsuda, M. J. 1989. Public response to racist speech: Considering the victim's story. *Michigan Law Review* 87, 2320-2381.

- Mattiello, E. 2022. Language aggression in English slang: The case of the -o suffix. In N. Knoblock (ed.) *The Grammar of Hate. Morphosyntactic Features of Hateful, Aggressive, and Dehumanising Discourse*. Cambridge: Cambridge University Press, 34-58.
- Moran, M. 1994. Talking about hate speech: Rhetorical analysis of American and Canadian approaches to the regulation of hate speech. *Wisconsin Law Review* 6, 1425–1514.
- Musolff, A. 2012. The study of metaphor as part of critical discourse analysis. *Critical Discourse Studies* 9(3), 301–310.
- Nuyts J., Byloo P. & Diepeveen, J. 2010. On deontic modality, directivity, and mood: The case of Dutch *mogen* and *moeten*. *Journal of Pragmatics* 42(1), 16-34.
- Oberschall, A. 2012. Propaganda, hate speech and mass killings. In P. Dojcinovic (ed.) *Propaganda, War Crimes Trials and International Law: From Speakers' Corner to War Crimes*. London: Routledge, 189-218.
- Over, H. 2021. Seven challenges for the dehumanization hypothesis. *Perspectives on Psychological Science* 16(1), 3-13.
- Paz María A., Montero-Díaz, J. & Moreno-Delgado, A. 2020. Hate speech: A systematized review. *Sage Open* 10(4), 1-12.
- Peterson, D. 2022. Homophobic space-times: Lexico-grammatical and discourse-semantic aspects of the softscapes of hate. In N. Knoblock (ed.) *The Grammar of Hate. Morphosyntactic Features of Hateful, Aggressive, and Dehumanising Discourse*. Cambridge: Cambridge University Press, 262-287.
- Poletto, F., Basile, V., Sanguinetti, M., Bosco, C. & Patti, V. 2021. Resources and benchmark corpora for hate speech detection: Asystematic review. *Language Resources and Evaluation* 55, 477-523.
- Ruzaitė, J. 2018. In search of hate speech in Lithuanian public discourse: A corpus-assisted analysis of online comments. *Lodz Papers in Pragmatics* 14(1), 93-116.
- Technau, B. 2018. Going beyond hate speech: The pragmatics of ethnic slur terms. *Lodz Papers in Pragmatics* 14(1), 25–43
- Tsesis, A. 2002. *Destructive Messages. How Hate Speech Paves the Way for Harmful Social Movements*. New York: New York University Press.
- UDHR: United Nations, General Assembly 1948. *Universal Declaration of Human Rights*. Department of State, United States of America.
- Waldron, J. 2012. *The Harm in Hate Speech*. Cambridge, Mass.: Harvard University Press.
- Ward, K. 1997. Free speech and the development of liberal virtues: An examination of the controversies involving flag-burning and hate speech. *University of Miami Law Review* 52(3), 733–792.
- Weaver, S. & Bradley, L. 2016. I haven't heard anything about religion whatsoever: Audience perceptions of anti-Muslim racism in Sacha Baron Cohen's *The Dictator*. *Humor* 29(2), 279-299.