

HATE SPEECH IN COMMENTS ON SOCIAL MEDIA: TOWARDS A CRITICAL DIGITAL CITIZENSHIP

Submetido em: 27/10/2024

Aceito em: 16/3/2025

Publicado em: 7/8/2025

Marlana Carla Peixoto Ribeiro¹

Hélvio Frank de Oliveira²

PRE-PROOF

(as accepted)

This is a preliminary, unedited version of a manuscript that was accepted for publication in Revista Contexto & Educação. As a service to our readers, we are making this initial version of the manuscript available, as accepted. The article will still be reviewed, formatted and approved by the authors before being published in its final form.

<https://doi.org/10.21527/2179-1309.2025.122.16612>

ABSTRACT

With internet access becoming increasingly democratic, social practices have gained new contours and daily adherence, building an inevitable “digital world”. Through communication processes produced on social networks, digital discursive performances operated with meanings and effects emerging from/in digital relationships between users. The present reflections are based on the empirical material generated in the study by Ribeiro (2023), which sought to identify and problematize, through online comments generated from posts by female celebrities on the social network TikTok, speeches that fell into the category “hate”, observing the socially produced meanings around digital practice. From the 96 comments on six videos posted on the pages of six female celebrities, the relationship between gender and race can be seen, which threatens white masculinity, permeating

¹ Faculdade Evangélica de Goianásia – Goianásia/GO, Brazil. <https://orcid.org/0000-0001-8096-2231>

² Universidade Estaduas de Goiás – UEG. Anápolis/GO, Brazil. <http://orcid.org/0000-0002-0553-8075>

HATE SPEECH IN COMMENTS ON SOCIAL MEDIA: TOWARDS A CRITICAL DIGITAL CITIZENSHIP

intersectionalities that cross ageism, LGBTphobia, misogyny, racism, sexism and fatphobia, or all of these at the same time or in some proportion. This guides a critical digital citizenship that promotes openness and a sense of plurality about what surrounds individuals to see and read digitally, as well as a critical scrutiny of what they will do, subsequently, with the human behind the screen.

Keywords: Digital performances; *TikTok*; Hateful comments; Online criticality.

DISCURSO DE ÓDIO EM COMENTÁRIOS NAS REDES SOCIAIS: RUMO A UMA CIDADANIA DIGITAL CRÍTICA

RESUMO

Com o acesso à internet cada vez mais democrático, as práticas sociais têm ganhado novos contornos e adesões diárias, construindo um inevitável “mundo digital”. Mediante processos de comunicação produzidos em redes sociais, as performances-discursivas digitais operam com sentidos e efeitos emergidos das/nas relações digitais entre usuários. As presentes reflexões partem do material empírico gerado no estudo de Ribeiro (2023), o qual buscou identificar e problematizar, por intermédio de comentários on-line gerados a partir de postagens de celebridades femininas na rede social TikTok, discursos que se inscreviam na categoria “ódio”, observando os sentidos socialmente produzidos em torno da prática digital. A partir dos 96 comentários em seis vídeos, postados nas páginas de seis celebridades femininas, nota-se relação entre gênero e raça, perpassando interseccionalidades que atravessam o etarismo, a LGBTfobia, a misoginia, o racismo, o sexismo e a gordofobia, ou tudo isso ao mesmo tempo ou em alguma proporção. Orienta-se, nesse contexto, a prática de uma cidadania digital crítica que promova abertura e senso de pluralidade sobre o que rodeia os indivíduos para ver e ler digitalmente, bem como um escrutínio crítico sobre o que o usuário faz com o humano que há por trás da tela.

Palavras-chave: Performances digitais; TikTok; Comentários de ódio; Crítica on-line.

HATE SPEECH IN COMMENTS ON SOCIAL MEDIA: TOWARDS A CRITICAL DIGITAL CITIZENSHIP

Introduction

An increasing number of digital technological developments in information and communication impact our ways of life (Reig; Vilchez, 2013). With access to the internet, social practices have gained new contours and daily adherence. Eminently digital, practices on computer screens, *smartphones*, etc., they begin to integrate our existence and build, in us, other ways of living in society. Founded on values and principles mobilized by *online relationships*, this entire dynamic is organized performatively and discursively, producing, circulating and consuming certain meanings (Blommaert, 2020; Butler, 2021).

Through communication processes produced on social networks, digital discursive performances materialize in textual forms, which operate with meanings and effects of meanings emerging from/in digital relationships between users. This condition is intersubjective and presents, as a problem, what we call digital hate, which makes it possible to verify, in the act of commenting *online*, several performative–discursive consequences of its promotion on social networks. Works such as those by Biar (2020), Costa (2022), Costa, Oliveira Junior and Frank (2021) and Ribeiro (2023) have already made some critical-discursive contributions to the field of critical applied linguistics. According to Louro (2015), although it is relevant at this social moment to censor, restrain or intimidate the propagators of hate speech, it is not enough to simply provide for it in law(s), with the purpose of punishing and criminalizing those who utter it. According to the author,

[...] whoever insults does not insult alone, but is, in fact, a speaker who echoes other voices. Censorship and the law can then target this speaker without being able to extinguish the speech that supports it and that continues to circulate more widely. [...] it would therefore be important to consider other ways of dealing with hate speech (Louro, 2015, p. 274).

To offer an educational and transdisciplinary option as a way of dealing with digital hate, in this text, we address “internet hate speech” as digital discursive practices produced in cyberculture, which materialize from a text, in this case, an *online* comment published on a social network, to problematize the implications of this performative-discursive act and its meaningful effects on social life, aiming for critical digital citizenship.

HATE SPEECH IN COMMENTS ON SOCIAL MEDIA: TOWARDS A CRITICAL DIGITAL CITIZENSHIP

In the next sections, we contextualize the study to discuss the relationship of social networks as discursive instances where texts are politically produced, from which comments that perform “hate speech” in digital discursive practices materialize, raising transdisciplinary issues around the object of study. We problematize, from a performative–discursive perspective, the effects of digital acts of hate, with their strength and perlocutionary effects. Finally, we reflect on its reverberations in digital discursive practices, raising the possibility of critical digital citizenship (Ribble, 2012).

1 Situating the study

Flannery (2021, p. 31) outlines the propositions manifested by hate speeches as those that “reveal[m] fixed and preconceived ideas of the other [...], [which are] marked[s] by positions constructed according to differences”. Trindade (2022, p. 17) also characterizes them as manifestations, verbal or written, of “thoughts, values and ideologies that aim to inferiorize, discredit and humiliate a person or social group, depending on characteristics such as gender, sexual orientation, religious affiliation, race, place of origin or class”.

Our reflections for this article are based on the empirical material generated in the study by Ribeiro (2023), which sought to identify and problematize, through *online comments* generated from video posts on the social network *TikTok* by female celebrities, discourses that are they entered the category of hate, observing the socially produced meanings around digital practice. First, the author searches *online*, opting for “celebrities who have suffered the most hate speech on the internet recently”, and finds the following celebrities as a result: Gretchen, Preta Gil, Luísa Sonza, Pablo Vittar, Karol Conká and Ludmilla. In the comparison carried out between July 2021 and February 2022, a first interesting finding emerges: most of them are women and people who are members of historically vulnerable groups owing to their bodies, identities and discursive performances. that deviate from the compulsory cisgender and heterosexual binary in force in Brazil.

From this, the author analyzes 96 comments generated on the pages of these six female celebrities after discursive performances that are posted and indicates the relationship between gender and race, which threatens white masculinity, permeating intersectionalities

HATE SPEECH IN COMMENTS ON SOCIAL MEDIA: TOWARDS A CRITICAL DIGITAL CITIZENSHIP

that cross ageism, LGBTphobia, misogyny, racism, sexism and fatphobia, or all of these at the same time or in some proportion. These findings prompted us to discuss, then, the linguistic act of *online commenting* on social networks, as well as the impact of this discursive performance on contemporary digital practice, because we understand that commenting *online* on social networks produces effects of meaning in digital/social relationships.

Our interest is to produce, from this, linguistic knowledge that helps people reflect and transform their own lives into everyday life (Borba, 2014; Moita Lopes, 2006, 2009). Furthermore, to interact digitally and ethically with other people, their differences should be respected. To achieve this, it is necessary to take a more attentive and sensitive look at the discursive lines of hate on social networks, pursuing meanings that go beyond the published comments, which may cause human suffering precisely because they reverberate inequalities, violence, exclusions, etc. (Butler, 2021).

1.1 From Web2 to social networks: digital discursive practices

Castells (1999) noted that the internet would be a means of (re)organizing social structures that would continually modify our human relationships. We can see cyberspace as a “space of human interaction”, provided by the global computer network (Lévy, 1999), as a new place for communication, organization and transaction, as well as a new information and knowledge market, in which praxis is “interactive, reciprocal, heterogeneous and “untotalizable” (Grillo, 2006, p. 17).

With the second generation of services offered by the internet, created since 2004, the so-called Web 2.0 or Web2ⁱ, countless interactions are carried out every day across the world at an unimaginable speed and reach. Web2 is configured as a locus for publishing, sharing and organizing information, expanding the interactive capacity among network users. The “era of connection”, mobility and instantaneity, as explained by Dias (2012), is marked by interaction and collaboration in the organization of content, encompassing countless languages and building a primordial aspect: digital practice, from which no one can tear themselves away. Once connected to the internet.

HATE SPEECH IN COMMENTS ON SOCIAL MEDIA: TOWARDS A CRITICAL DIGITAL CITIZENSHIP

As it starts to promote relationships between its networked users, Web2 allows digital sharing and transmission, from anywhere, in different ways, to different people, in a kind of “technological web” (Borges, 2020, p. 35). This reveals the relevance of linguistic concerns in social constructions, which are consequently digital, because critical thinking is not always stimulated in search and processing or in the interaction and sharing of information to generate knowledge and ethical principles in the relationships between users in digital environments.

Faced with this insatiable and often inconsequential handling of technological-digital devices, we cannot neglect the discursive problem of social networks in contemporary times located in the performances of users who type texts in their digital communication relationships. We therefore highlight the relationship, involvement and attention to this performative–discursive character manifested on digital screens, with meanings and effects of meanings that produce realities.

1.2 Social networks and the discursive problem

The emergence of social networks gave the user society a new way of life: new forms of social interaction and, consequently, new “ways of being in the world” (Lapa et al., 2020, p. 40). As the authors suggest, physical spaces for human coexistence, limited by time and geographic location, have transcended into digital spaces, in which many people have begun to express themselves, thus assuming the famous idea of Amossy (2014) that social networks have become the public square of the 21st century (Cabral, 2019).

Effectively, social networks have provoked a new digital and dialogical sociability, seen as “a space of free access, where people can debate and develop arguments about issues of common life” (Touraine, 1994 apud Lapa et al., 2020, p. 28). From digital life, individual and collective behaviors, values are (re)affirmed, (re)producing digital meanings, which are translated socially. For Nonato and Sales (2019, p. 150), social networks have acted as “mediators of man’s being and acting in the world”, changing the way in which he establishes relationships with others, with space and with knowledge, adding different languages and techniques and enabling the experience of varied creative processes

HATE SPEECH IN COMMENTS ON SOCIAL MEDIA: TOWARDS A CRITICAL DIGITAL CITIZENSHIP

(Hetkowski; Menezes, 2019).

Through social networks, “any individual with access to the internet is able to produce entertainment, educational or opinion content and share it so that other people can see it” (Nandi, 2018, p. 13), the discursive characteristic of the digital sphere, with its uses and human relationships, deserves reflection. Along with digital practices guided by the perspectives of relating, influencing each other, and expressing a point of view, approximations or divergences between users will often (re)construct meanings based on shares, likes and comments online.

In contemporary times, social networks have become the main daily communication channel for most people. For example, users express themselves through publications on their profiles, thoughts and feelings, and in a dialogical-like environment, discursive spaces are produced for conversations and positions, among other attitudes mobilized from manifestations between users. This entire way of reading, producing and circulating texts (Rojo, 2013) generates negotiations of meaning, discursively promoting debates and statements, among other disputes, depending on the users and their relationships established on the network.

As dynamic systems, which come from “an environment of interaction between culture, society and information technology” (Santaella; Cardoso, 2014, p. 74), digital discursive practices democratize global communication and, however, signal precisely what society (consciously or unconsciously) promotes. The digital here is a result or circumstance of the social structure that is already forged between the networks and the streets (Beiguelman, 2019) under intentional discursive action. The “like” button, for example, for Odriozola (2020), acts as a positive feedback effect, connecting to a user’s sense of belonging or even (re)knowledge, whereas the intentionality of those who produce the post can perform a “look” you want to build. As highlighted by Lapa et al. (2020), in the routine of exposing life in real time, with a potential and ephemeral character, but which, at least apparently, allows some other person's point of view as a return (Nonato; Sales, 2019).

Amid the multiple linguistic and semiotic modalities provided by social networks, meanings are mobilized through the expression of bodies, shapes, colors, effects, movements,

HATE SPEECH IN COMMENTS ON SOCIAL MEDIA: TOWARDS A CRITICAL DIGITAL CITIZENSHIP

gestures, performances, thoughts, and sounds, among other (effects of) meanings produced as “publication” or “post”, which, in the relationship between users, with the intention of creating or not creating social ties, generate disputes, negotiations and/or demarcations of meanings foreseen in a like, a share or even a comment. In the digital discursive dynamics on social networks, online comments, for performative purposes, can materialize from writing to a like click, as this can also signal some type of endorsement of the published text.

The interactive characteristic of users on social networks is elementary to the performative-discursive tone of the act of commenting (Austin, 1990), making the relationship between users even more complex, owing to enunciative-pragmatic aspects, such as the intentionality of production (the content creator wishes to cause in the follower to comment on the post), the conditions of production, circulation and consumption of the message (interlocution relationships between users that guarantee the digital enunciative character). The meanings conveyed in the comments themselves (the materiality of the content with the illocutionary force of the message itself and the perlocutionary effects as a condition that drives the interlocutor to interpret and comment), and the social network as an instance of discursive production in contemporary life, producing identities and differences, since

[the] individual becomes an active agent within the internet, creating and shaping their relationships and customizing their profiles according to their interests and the audience they intend to reach. Furthermore, the two-way communication provided by the change in the sending pole allows the receiver to send feedback to the sender, thus asserting itself as a less passive part of the communication process (Borges, 2020, p. 39).

Although the complexity of commenting on social networks becomes interactive and expands according to the availability of resources—insertion of (hyper)links and mobilization of different discursive genres, for example—the central point is that digital praxis materializes discursively based on “manufactured” comments around a publication, establishing relationships among creators, content and digital readers. For this reason, they deserve critical scrutiny to problematize their circumstances, since from the digital culture that is spreading, a symbiotic relationship emerges, with mutual and cooperative influence,

HATE SPEECH IN COMMENTS ON SOCIAL MEDIA: TOWARDS A CRITICAL DIGITAL CITIZENSHIP

between society, culture and new technologies.

The digital condition elevates the human issue, through the discursive aspect, by providing the user with the possibility of becoming whatever they want, evoking poststructuralist constructions of language. Similarly, the internet algorithm is intertwined with the conversational factor, allowing the triggering of linguistic manifestations of a typed reference, causing viralization, popularization or even monetization by moving certain content and themes. In search of interaction, sometimes random, algorithms almost always end up operating linguistically on promoted digital values of influence and visibility.

1.3 Online comment (text) and its flirtation with 'hate speech' (speech)

Online commentary is a discursive textual materiality foreseen in certain genres of the digital sphere, such as social networks (Costa, 2022; Costa; Oliveira Junior; Frank, 2021; Ribeiro, 2023). Generally, it starts from an interactive reception by those who read a publication as source text and is the result of a digital practice, through which hate can be discursively conveyed. This occurs because any publication made on a social network, as an environment for relationships between users, for the exhibition of ideas, presupposes negotiation and dispute. A comment on social networks is, therefore, an opinion that materializes in the appreciation and interpretation of a first text posted by another internet user.

The comment, in general, combines multimodal linguistic and semiotic elements, reflecting the issuance of an opinion/evaluation position or reaction toward the digital that was previously viewed. Social networks such as *TikTok*, *Instagram* and *Facebook* are short, condensed and informal, with the use of abbreviations and icons and the absence of punctuation and concern with the culture norm of the written Portuguese language (Grillo, 2006). It is also common to use *emoticons* and *hashtags*, which function as resources for aggregating data about a certain subject (Seixas, 2021). Furthermore, an *online* comment on a social network, motivated by an initial post, will always be a position in dispute, through which there may be positive or negative receptions, representing (dis)identifications. This condition is dynamic to the point of motivating everything from likes, shares, and comments

HATE SPEECH IN COMMENTS ON SOCIAL MEDIA: TOWARDS A CRITICAL DIGITAL CITIZENSHIP

to digital lynching/cancellation.

There is a digital cycle of discursive action arising from the comment, which does not end, as it moves from the discursive performance of those who produce the motivating text to generate the comment to sociohistorical, cultural and political aspects transcendental to the text, which culminates in the senses and effects of reader senses that will be mobilized and, eventually, put into confrontation on the basis of the publication, the comment and/or the commentary. From this perspective, dynamics can produce other digital genres, such as a type of debate, which reflects the need of the internet user-reader to publish his opinion as a position in relation to what he saw posted. In this case, the comment reproduces a position taken on a post or even on another comment (Costa, 2022; Costa; Oliveira Junior; Frank, 2021; Ribeiro, 2023).

Once it becomes a performative utterance, what is published on a social network, specifically as a comment, starts to produce meanings and effects of meaning among users of that social network. It is at this moment that comments appear as discursive achievements that involve self-care based on a posted digital imagery arrangement. In this sense, the source text that generates the comment can build identity networks, forming a kind of digital bubble between users, or even a refusal to accept a read position.

Importantly, with the help of algorithms and resources specific to computer language, everything that is written and commented on social networks serves as a database that, in the future, will be used to build affinities between the content that the user of a social network will have at your disposal, even if involuntarily, will be part of your digital world and that will certainly translate into your thematic *online browsing engagements*.

Although this issue is not a prominent point in our discussion, it is important not to escape criticism of the algorithm because it also provides the provision of thematic repertoires based on what we search for *online*, generating identifications, belongings, and curiosities of those who use and participate in certain topics. In this contradictory fluidity between posts and comments, which is surrounded by capitalist reasons, it is necessary to educate ourselves linguistically that, behind a comment, there will be hateful motivations that range from identification and nonidentification with the meanings conveyed, even the

HATE SPEECH IN COMMENTS ON SOCIAL MEDIA: TOWARDS A CRITICAL DIGITAL CITIZENSHIP

intention of generating something profitable with the comment: the popularity of the commenter or likes through the number of *likes*, shares or even responses to comments, which are sometimes completely impactful under the radical purpose of generating controversy.

In this sense, we emphasize that commenting is a discursive-performative act (Austin, 1990), as it does not involve a mere exchange of messages between users. In contrast, with practice, spaces built within and by difference are forged, which translates into constant negotiations and disputes over “digits”, meanings and discursive mobilizations. Our studies observing *online* comments on social networks (Costa; Oliveira Junior; Frank, 2021; Ribeiro, 2023) revealed that the discursive practice of digital hate starts from comments that generate exclusion, denial or even abjection of certain bodies, identities and performances that, for whatever reason, escape social expectations and demands under white, heterosexual and cisgender conditions.

In this dimension, the comment is an idea that can be discursively guided by hatred, based on issues that generally involve the difference observed in the other, arising from the way in which human beings are educated to address differences. Intensified with the digital advent, hate—which, according to Mandela (1994), accompanies humanity throughout its history—has expanded as an increasing number of users have joined social networks. For Brugger (2007), so-called *haters* use offenses, insults, intimidation and harassment generally related to identity issues of race, color, body, ethnicity, nationality, social class, sex, sexuality, religion, and other dissident intersectionalities on display. They therefore have the capacity to instigate violence, hatred or discrimination against certain people, cultures, bodies and performances, aiming to target users and groups, causing human suffering, exclusion, and violence.

The fact is that social networks have recently become the stage for episodes of hate promulgation produced by *online comments*, often mobilized by aggressive, offensive writing and with subtext marked by the disqualification of the other person/group that is produced or appears in the text. Instead of being guided by the defense of a different and divergent point of view in the face of the typed, hate speech generally has as a strategy,

HATE SPEECH IN COMMENTS ON SOCIAL MEDIA: TOWARDS A CRITICAL DIGITAL CITIZENSHIP

according to Oliveira (2018), Guercio, Matheus and Napolini (2021) and Sarmento (2006), the expression of ideas of contempt or intolerance against certain people or groups, belittling them, disqualifying them or making them inferior simply because they belong to a certain community, with motivations linked to ethnic, religious, gender intolerance, physical appearance, sexual orientation, and nationality, among others.

Hate speech therefore concerns the devaluation of the other person and the failure to recognize the value in what the other person can offer. This can lead to the inferiorization and even dehumanization of people/groups (Dias; Ripoll, 2021), mobilizing a vast and complex repertoire of racist, sexist, misogynistic, elitist and homotransphobic statements and oppositions. Discursive studies on digital hate are not rare (Dias; Ripoll, 2021; França, 2019; Grillo, 2006), even in applied linguistics (Costa, 2022; Costa; Oliveira Junior; Frank, 2021; Ribeiro, 2023).

The common fact is that hateful comments are directed at bodies that contradict cisgender heteronormativity. In online searches, it is possible to see that most celebrities targeted by hate speech are women or recognize themselves from a “deviant” sexual orientation or gender identity (Guercio; Matheus; Napolini, 2021) or an ethnic identity other than whiteness. These people, as they publish their digital performances, suffer digital acts of violence, based on hateful comment texts, vulgarization, inferiorization and the claim of a heterocisnormative relationship as a universal condition.

As an act of language, the published comment, considered “from whom to whom”, refers to the possibility of language conveying hatred based on what it supposedly describes (Borba, 2014; Pinto, 2007; Urzêda-Freitas, 2020). Paradoxically, digital discursive hate, owing to the algorithmic techniques of social media and the *online complexity* that surrounds it, instead of being refuted, can curiously be projected and throw a user or circumstance of hate into the spotlight:

[...] Another characteristic of digital platforms that can contribute to the circulation of hate speech is their operating logic, which is based on algorithms whose management of content visibility is directly related to user preferences (GILLESPIE, 2018) – this is what some authors call it 'echo -chambers'. By prioritizing content in timelines depending on the individual's opinion, algorithms

HATE SPEECH IN COMMENTS ON SOCIAL MEDIA: TOWARDS A CRITICAL DIGITAL CITIZENSHIP

can also produce a distorted perception of social scenarios, in addition to inducing polarization through the creation of ideological bubbles (Ruediger; Grassi, 2021, p. 21).

The problem occurs when the propagation of hate speech denies its victims the right to equality, which is inseparable from freedom (Cintra, 2012). This is because digital hate acts to undermine access to the possibility of strangeness when faced with socioidentity differences and by nullifying or silencing people who are already marginalized in society. Thus, more than a contrary thought, what is submerged is the attempt to reject what belongs to the other, triggering detached praxis of reflexivity about one's own doing *online* and enabling a lack of ethical commitment and digital violence.

Nandi (2018) stated that when a person addresses another person based on hate speech, they violate dignity in its intersubjective dimension in the respect that each human being owes to the other. Moreover, in addition to direct attack, the dignity of people who recognize themselves as having those characteristics that give rise to violence is also attacked. This is the so-called “diffuse victimization” in which countless victims who identify with the target object of the attack can be affected, even if they are not directly related to the event or the direct victim.

Digital hate constitutes an attitude of delegitimization of others, which can generate the attribution of less dignity and denial of corresponding equal respect, in addition to the ruin of self-esteem in society, with a view to destroying the positive evaluation of one's character and history shared (Martins, 2019). With the aim of silencing the “divergent”, placing them on the margins of any discussion, hatred becomes a discursive strategy, often articulated, that aims to silence a part of the population, mitigating their strengths and generating power for the offender (Guercio; Matheus; Napolini, 2021). In the words of the authors,

[the] agent or the one we call the precursor of hate speech, completely denies the recognition of the other, the one we call the victim or target of aggression, as being someone incapable of possessing characteristics that link him to a dignified universal position, in addition to deny or simply not accept their identity attributes (Guercio; Matheus; Napolini, 2021, p. 45).

HATE SPEECH IN COMMENTS ON SOCIAL MEDIA: TOWARDS A CRITICAL DIGITAL CITIZENSHIP

For these reasons, the hate found on social media, which we call digital hate, is fundamentally discursive. Therefore, it should not only be repelled and denied but also understood from the perspective of alerting to its potential meaning effects as a digital practice. This is relevant because hate speech often threatens, disqualifies and attacks diverse opinions, bodies (gender, sexuality, race, ethnicity), identities, behaviors, cultures, etc., under an intentional and linguistically defined sayable.

Amid the circumstances of what is said, what is typed becomes discursive, pursuing colors and bodies that enact transgression to social norms, performing differences. From this perspective, hate speech can be conceived, therefore, as modes of action—the emission (posting) and its reception (comments), ways of acting on oneself and in the world and, especially, on other people—thereby manifesting, in this way, the human capacity to (re)produce relations of domination for the (re)constitution, directly or indirectly, of all social structures.

3 Critical digital citizenship

Ribeiro (2023) developed research in which the meeting point of her empirical material led her to very intriguing conditions of analysis regarding celebrities who were in the *ranking* of hate speeches in the Brazilian context. The results therefore open space for us to delve deeper into issues of identity and differences that are linked to the “hate speech” suffered not only by famous people but also by ordinary people, as we have constantly monitored. The comment that reveals/incites hatred is a challenge associated with the marginalized existential corporality and, mainly, materialized in discursive performances and in the corporealities of difference.

Digital hate weighs on so-called feminine and/or dissident bodies that deviate from modern binarism. In a white-heteronormative world, there is no denying colonizing principles that reject those people who transgress certain norms. Other dimensions, such as ageism and fatphobia, are included in the lists of intersectionalities that help to proportionate this spread. Simply by existing and speaking in the digital world, these bodies of difference will be attacked with violent comments, trying to stop discursive performances that escape

HATE SPEECH IN COMMENTS ON SOCIAL MEDIA: TOWARDS A CRITICAL DIGITAL CITIZENSHIP

the conventional social order.

In a plural society, however, we must reflect on comments on social networks to build a less violent digital world with less human suffering. To this end, we propose critical digital citizenship in applied linguistics, which moves between other areas of knowledge, to the act of commenting on social networks, considering the production and discursive effects of this practice in the digital world. If the comments and the entire digital dynamic start from a discursive-performative characteristic, the idea is to raise awareness in society so that lives are not “cancelled”.

We emphasize that digital hate cannot be considered mere opinion and is not supported by freedom of expression. In addition to enacting laws, we can promote the conscious use of social media. For this purpose, critical digital citizenship would then be a strategy for resignifying *online praxis* along a path of emancipation and autonomy in which discursive practices, digitality and education come together in search of critical sense, considering and understanding respect social diversity for a more peaceful and human coexistence *online* and to confront injustices (Alencar, 2022).

In addition to the difficulty of criminalizing and punishing those who spread hateful speeches, especially given the dispersion of meanings promulgated *online*, as they can operate as a subjective interpretation of those who do/receive it, the ethical and critical use of TDIC is referenced in educational proposals guided by documents such as the Common National Curricular Base (BNCC). It envisages school reflection on the use of media, involving aspects of citizenship, as children and young people have appropriated various digital tools, to the point of spending most of their time connected to the internet.

As digital users, it is beneficial for us to observe what we are mobilizing on the networks when we come across what makes us different, in a permanent effort to critically perceive how we are being in the world (Freire, 2005). Social networks should be environments built to discuss choices, conduct, and political and identity positions, always striving for respect for what goes against our way of thinking. Fluid spaces of (mis)learning, temporary knowledge and uncertainties should be created, with a view to exercising self-awareness, self-criticism and the ability to dialog with respect and empathy. In a complex

HATE SPEECH IN COMMENTS ON SOCIAL MEDIA: TOWARDS A CRITICAL DIGITAL CITIZENSHIP

space of discursive performances, with a multiplicity of meanings, the negotiation of difference would be necessary for good digital coexistence.

It is part of digital citizenship, more than knowing how to use technological tools, a more attentive and sensitive look at the relationships that take place there and the way in which the meanings in these achievements are interpreted. To this end, attention must be paid to the discursive to observe power relations that circulate in social instances, especially on the internet (Pereira, 2018), so that, from this, people develop “the critical capacity to perceive and question values that are conveyed through institutionally legitimized discourses and power relations that are constantly presented as 'natural' and, therefore, not subject to criticism and questioning” (Pereira, 2018, p. 53).

From the act of commenting on social media, we can shame, humiliate, colonize, in short, act in the world through violence, or the opposite: we can cultivate difference and respect. Therefore, it is necessary to have discursive and digital awareness to check information, question and problematize (one's own) reality, adopting observatory and contesting attitudes in relation to dominant and oppressive discourses, with a view to promoting social changes (Pereira, 2018). Unlike consensus and harmony, social networks need to retain spaces of difference, dialog, relationships, and affective resistance. In a multicultural digital society (Guercio; Matheus; Napolini, 2021) and constant estrangement, peaceful coexistence can be guaranteed by respect to users in an environment that does not trivialize the existence of the human behind the screen, suggesting that content posts and user comments promote meanings capable of promoting the expansion of thought and leaving our own epistemic place (Domingues; Paiva; Rosa-da-Silva, 2021; Guercio; Matheus; Napolini, 2021).

Vulnerable people and groups have the socio-historical right to digital existence, especially because this can promote the visibility of differences and relevant changes in society. In this sense, it is important to confront, in a critical and ethical way, the comments that fall into the discursive praxis of digital hate, especially those directed at minoritized groups and peopleⁱⁱ, to confront social inequalities and injustices. It is therefore pertinent that digital spaces are constructed as places of exchange, dialog, listening, will, reception,

HATE SPEECH IN COMMENTS ON SOCIAL MEDIA: TOWARDS A CRITICAL DIGITAL CITIZENSHIP

acceptance and mutual solidarity. However, it would be relevant, at the same time, to have critical engagement in problematizing what is read and disseminated on social networks.

In reference to Butler (2021), digital criticism can start from questioning who types the standard, who is considered when a convention speaks, and at what moment this convention speaks. Such questions even provide us with clues for interpreting/understanding the text, something very relevant to contemporary times and can provide urgent digital citizenship in a space that is intended for dialog, sharing experiences and exposing thoughts, values, culture, life, etc., but above all, it can make us think, expand our own perception of ourselves, culminating in the malleability of our own certainties and truths around established social values about gender, race, and other differences. The idea is that, through the sharing of opinions and thoughts, communication, respect, empathy and the exercise of otherness on social networks are encouraged, including providing opportunities for intercultural interaction and the promotion of social equality on digital networks.

Finally, we need openness and a sense of plurality about what surrounds us to see and read digitally, as well as critical scrutiny of what we will do next, with the human behind the screen. Moreover, it is necessary to act against prejudiced and discriminatory digital attitudes in these environments. Questioning other people's truths can be an incentive to understand and value recognition of the plurality that presents itself digitally. Are we, as Costa, Oliveira Junior and Frank (2021) assert, compromising and canceling lives or contributing to a sensitive world and in favor of difference?

Hate speech, as a phenomenon of multiple origins and natures, on the other hand, also demands efforts, such as norms and judicialization on the use of the internet and social networks, seeking a civility capable of reflecting subjects, bodies and groups historically marginalized, welcoming and demystifying the multiplicity that constitutes us as human beings, especially through education.

Although social media invokes the right to speak and give an opinion without restrictions, this is not legally or ethically defensible. Hate speeches, as they reject the social, historically and culturally constructed plurality in Brazil, often translate into insults and offenses of a racial, misogynistic, LGBT-phobic, ageist, fat-phobic, and xenophobic nature,

HATE SPEECH IN COMMENTS ON SOCIAL MEDIA: TOWARDS A CRITICAL DIGITAL CITIZENSHIP

among other discriminatory speeches in digital communication (Flannery, 2021), almost always with the intention of silencing and weakening the people and/or groups attacked (Guercio; Matheus; Napolini, 2021). Digital discursive praxis in favor of difference can act to provide alternative and possibly more effective approaches to the conscious use of digitality, contributing to a better understanding of the breadth and potential for harm of what is conveyed on the networks, promoting more in-depth debates and consistency about hatred, and continually reinforcing the importance of reflecting the origin and purpose of intolerant behaviors and adherence to ideas that adapt to the use of violence.

Final considerations

Castells (1999) shows that if the internet has reorganized social structures, an ethical look at the issue can result in what tool can provide us with the collective construction of projects that change society. Nonato and Sales (2019, p. 150), invoking Castells (2002), noted that “new technologies are transforming the way we produce, consume, manage, live and die, certainly not for themselves, but as powerful mediators of a larger set of factors that determine human behavior and social organization”. In this sense, digital culture, far from being anchored in neutrality, promotes the circularity of subjectivities, intentions and ideological positions, which need to be discussed and analyzed from the perspective of education (Melo; Silva, 2018).

Cabral (2019) noted that when people comment on social networks, as a public act, they need to be aware that their behavior may be examined, implicitly accepting the consequences of choosing to communicate in an environment accessible to countless other people. Living in this “postdigital society” (Blommaert, 2020), which is hypercomplex and enhances hypersociability (Santaella, 2013), also requires “the ability to distinguish between different languages and media, their specific communicative natures, their political-social injunctions and, from this, being able to develop the ability to raise questions about everything we read, see and hear” (Gregolin, 2017, p. 171). Thus, the user of social networks needs, in addition to dealing with technical issues, also to be able to perceive the matrix of power that structures society (and, consequently, the network society), reflecting on the

HATE SPEECH IN COMMENTS ON SOCIAL MEDIA: TOWARDS A CRITICAL DIGITAL CITIZENSHIP

processes of inclusion and exclusion that surround the internet and how everyone participates in it, often even thoughtlessly.

The fluidity and instability of digital discursive practices denote the “importance of a path of collective awareness” (Kilomba, 2019, p. 12), considering the fundamental notion that the broader and more complex structuring of consciousness is a continuous and dynamic process of humanization and liberation (Freire, 2005), of overcoming injustice and other forms of oppression and cruelty.

REFERENCES

ALENCAR, Claudiana Nogueira de. “Deixar existir a poesia mais selvagem” na tessitura de uma gramática anticolonial. In: LANDULFO, Cristiane; MATOS, Doris. (org.). *Suleando conceitos em linguagens: decolonialidades e epistemologias outras*. Campinas: Pontes, 2022. p. 09-12.

AMOSSY, Ruth. *Apologie de la polémique*. Paris: Presses Universitaires de France, 2014.

AUSTIN, John L. *Quando dizer é fazer*. Trad. Danilo Marcondes de Souza Filho. Porto Alegre: Artes Médias, 1990.

BEIGUELMAN, Giselle. Redes reais: arte e ativismo na era da vigilância compartilhada. *Rapsódia*, São Paulo, v. 12, p. 65-78, 2019. Disponível em: <https://bit.ly/3jBNtWc>. Acesso em: 29 maio 2022.

BIAR, Liana de Andrade. “(Não) Leia os comentários”: a disputa da notícia sobre o assassinato de Marielle Franco. *Trabalhos em Linguística Aplicada*, Campinas, v. 59, n. 2, p. 1051-1069, 2020.

BLOMMAERT, Jan. O discurso político em sociedades pós-digitais. *Trabalhos em Linguística Aplicada*, Campinas, v. 59, n. 1, p. 390-403, 2020. Disponível em: <https://bit.ly/3PY5wSc>. Acesso em: 17 set. 2022.

BORBA, Rodrigo. A linguagem importa? Sobre performance, performatividade e peregrinações conceituais. *Cad. Pagu*, Campinas, n. 43, p. 441-474, dez. 2014. Disponível em: <https://bit.ly/3QINnqz>. Acesso em: 20 nov. 2020.

BORGES, Clara Andrade. *Fast-Foodização da informação: a comunicação em tempos de cibercultura e modernidade líquida*. 2020. 82 f. Monografia de graduação (Bacharelado em Mídias Digitais) – Centro de Ciências Humanas, Letras e Artes, da Universidade Federal da

**HATE SPEECH IN COMMENTS ON SOCIAL MEDIA:
TOWARDS A CRITICAL DIGITAL CITIZENSHIP**

Paraíba, João Pessoa, 2020.

BRASIL. Ministério da Educação. *Base Nacional Comum Curricular*. Brasília, 2018.

BRUGGER, Winfried. Proibição ou proteção do discurso do ódio? Algumas observações sobre o direito alemão e o americano. Trad. Maria Ângela Jardim de Santa Cruz Oliveira. *Revista de Direito Público*, ano 4, v. 15, n. 117, jan./mar, p. 117-136. 2007. Disponível em: <https://bit.ly/3gxEnsp> . Acesso em: 26 nov. 2022.

BUTLER, Judith. *Discurso de ódio: Uma política do performativo*. Trad. Roberta Fabbri Viscardi. São Paulo: Editora Unesp, 2021.

CABRAL, Ana Lúcia Tinoco. Violência verbal e argumentação nas redes sociais: comentários no Facebook. *Calidoscópio*, Porto Alegre, v. 17, n. 3, p. 416-432, set./nov. 2019.

CASTELLS, Manuel. *A sociedade em rede*. São Paulo: Paz e Terra, 1999. V. 1.

CINTRA, Reinaldo Silva. *O discurso de ódio sob uma teoria performativa da linguagem*. 2012. 92 f. Monografia (Bacharelado em Direito) – Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2012.

COSTA, Diêgo Martins da. *Análise discursiva crítica de comentários on-line motivados por publicações de notícias relacionadas à violência contra mulheres em um perfil jornalístico do Instagram*. 2022. 138 f. Dissertação (Mestrado em Educação, Linguagem e Tecnologias) – Universidade Estadual de Goiás, Anápolis, 2022.

COSTA, Diêgo Martins da; OLIVEIRA JUNIOR, Paulo Almeida de; FRANK, Hêlvio. Comentários em uma página da rede social Instagram: reflexões situadas de uma prática discursiva on-line. *Muiraquitã: Revista de Letras e Humanidades*, Rio Branco, v. 9, n. 2, p. 58-74, jun./dez, 2021. Disponível em: <https://bit.ly/3jy2EzB>. Acesso em: 25 jun. 2022.

DIAS, Kamyla Stanieski; RIPOLL, Daniela. Discurso de ódio e representações de corpos de mulheres gordas no YouTube: uma análise dos canais “Tá querida” e “Alexandrismos”. *Revista diversidade e educação*, Rio Grande, v. 9, n. 2, p. 147-176, 2021.

DIAS, Stéphane Rodrigues. *A linguística na cultura digital: uma reflexão nas interfaces*. 2012. 86 f. Dissertação (Mestrado em Letras) - Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2012.

DOMINGUES, Rosania Gomes da Silva; PAIVA, Joana de Oliveira Ferreira. ROSA-DA-SILVA, Valéria. Educação para a diversidade pelo viés do letramento crítico. In: ROSA LOPES, Cristiane; CONTI DE FREITAS, Carla; AVELAR, Michely Gomes (org.). *Linguagens em tempos inéditos: desafios praxiológicos da formação de professoras/es de línguas: volume dois*. Goiânia: Scotti, 2021. p. 115-134.

**HATE SPEECH IN COMMENTS ON SOCIAL MEDIA:
TOWARDS A CRITICAL DIGITAL CITIZENSHIP**

FLANNERY, Mércia Regina Santana. *Nós versus eles: discurso discriminatório, preconceito e linguagem agressiva na comunicação digital no Brasil*. São Paulo: Editora CEPE, 2021.

FRANÇA, Thiago Alves. *Sentidos e funcionamentos do discurso de ódio em espaços do Facebook: uma leitura discursiva*. 2019. 276 f. Tese (Doutorado em Letras) – Universidade Federal de Pernambuco, Recife, 2019. Disponível em: <https://bit.ly/3vkjrst>. Acesso em: 23 nov. 2022.

FREIRE, Paulo. *Pedagogia do Oprimido*. 42. ed. Rio de Janeiro: Paz e Terra, 2005.

GREGOLIN, Isadora Valencise. Impactos das redes sociais nas práticas linguísticas: (re)pensando atividades para o ensino de estratégias de cortesia em língua espanhola. *Caracol*, São Paulo, n. 13, p. 163-178, jan./jun., 2017.

GRILLO, Maria Helena Corrêa Nogueira. *Blogbeca801: Um evento textual discursivo*. 2006. 216 f. Dissertação (Mestrado em Linguística) – Universidade Federal de Santa Catarina, Florianópolis, 2006.

GUERCIO, Cilene Rebelo Nogueira; MATHEUS, Rosemeire Solidade da Silva; NASPOLINI, Samyra Haydêe Dal Farra. A internet e o discurso de ódio: uma análise teórico-social. *Revista de Sociologia, Antropologia e Cultura Jurídica*, v. 7, n. 1, p. 42-59, jan./jul., 2021.

HETKOWSKI, Tânia Maria; MENEZES, Catia Nery. Práticas de multiletramentos e tecnologias digitais: múltiplas aprendizagens potencializadas pelas tecnologias digitais. In: FERRAZ, Obdália (org.). *Educação, (multi)letramento e tecnologias: tecendo redes de conhecimento sobre letramentos, cultura digital, ensino e aprendizagem na cibercultura*. Salvador: EDUFBA, 2019. p. 205-230.

KILOMBA, Grada. *Memórias da plantação: episódios de racismo cotidiano*. Trad. Jess Oliveira. Rio de Janeiro: Cobogó, 2019.

LAPA, Andrea Brandão; COELHO, Isabel Colucci; SCHWERTL, Simone Leal; LACERDA, Andreson Lopes de. *Redes sociais como espaço alternativo para formação crítica*. 2020. p. 21-43. Disponível em: <https://bit.ly/3VtmWY0>. Acesso em: 05 maio 2022.

LÉVY, Pierre. *Cibercultura*. 3. ed. São Paulo: Editora 34, 1999.

LOURO, Guacira Lopes. Discursos de ódio. In: SEFFNER, Fernando; CAETANO, Marcio. (org.). *Cenas Latinoamericanas da diversidade sexual e de gênero: práticas, pedagogias e políticas públicas*. Rio Grande: Ed. da FURG, 2015. p. 268-279.

MANDELA, Nelson Rolihlahla. *Long Walk to Freedom*. Boston: Edition Little, Brown & Co, 1994.

**HATE SPEECH IN COMMENTS ON SOCIAL MEDIA:
TOWARDS A CRITICAL DIGITAL CITIZENSHIP**

MARTINS, Anna Clara Lehmann. Discurso de ódio em redes sociais e reconhecimento do outro: o caso M. *Revista Direito GV*, São Paulo, v. 15, n. 01, p. 01-30, jan./abr. 2019.

MELO, Keite Silva de; SILVA, Andréa Villela Mafra da. Desafios e possibilidades da cibercultura para a educação. *Revista de Letras da Rural*, Seropédica, v. 3, p. 16-35. 2018.

MOITA LOPES, Luis Paulo da. (org.). *Por uma Linguística Aplicada Indisciplinar*. São Paulo: Parábola, 2006.

MOITA LOPES, Luis Paulo da. Linguística Aplicada como lugar de construir verdades contingentes: sexualidades, ética e política. *Gragoatá*, Niterói, n. 27, v. 2, p. 33-50. 2009. Disponível em: <https://bit.ly/3WL15fN>. Acesso em: 30 nov. 2021.

NANDI, José Adelmo Becker. *O combate ao discurso de ódio nas redes sociais*. 2018. 58 f. Trabalho de Conclusão de Curso (Bacharelado em Tecnologias da Informação e Comunicação) – Universidade Federal de Santa Catarina, Araranguá, 2018.

NONATO, Emanuel do Rosário Santos; SALES, Mary Valda Souza. Educação e os caminhos da escrita na cultura digital. In: FERRAZ, Obdália. (org.). *Educação, (multi)letramento e tecnologias: tecendo redes de conhecimento sobre letramentos, cultura digital, ensino e aprendizagem na cibercultura*. Salvador: EDUFBA, 2019. p. 141-172.

ODRIOZOLA, Amaia. Vício em redes sociais dispara na pandemia, mas há como recuperar o controle e se desintoxicar. *El País*, publicado em 12 out. 2020. Disponível em: <https://bit.ly/3WNRm8M>. Acesso em: 15 ago. 2022.

OLIVEIRA, Raphael Rodrigues Cardoso. *Discurso de ódio no Brasil: a complacência do Estado Democrático de Direito com o uso da liberdade de expressão como violência ao Outro*. 2018. 123 f. Monografia (Bacharelado em Direito) – Universidade de Brasília, Brasília, 2018.

PEREIRA, Ariovaldo Lopes. Construção de criticidade em espaços de atuação religiosa, política e acadêmica. In: PESSOA, Rosane Rocha; SILVESTRE, Viviane Pires Viana; MONTE MOR, Walkyria (org.). *Perspectivas críticas de educação linguística no Brasil: trajetórias e práticas de professoras/es universitárias/os de inglês*. 1. ed. São Paulo: Pá de Palavra, 2018. p. 47-57.

PINTO, Joana Plaza. Conexões teóricas entre performatividade, corpo e identidades. *D.E.L.T.A*, São Paulo, v. 23, n. 1, p. 01-26, 2007. Disponível em: <https://bit.ly/3I5BFp2>. Acesso em: 17 out. 2021.

REIG, D.; VÍLCHEZ, L. F. *Jovens na era da hiperconectividade: tendências, chaves e perspectivas*. Madrid: Fundacion Telefonica e Fundación Encuentro, 2013.

RIBBLE, Mike. *Digital Citizenship in Schools*. 3. ed. Nova York: Iste publishing, 2012.

**HATE SPEECH IN COMMENTS ON SOCIAL MEDIA:
TOWARDS A CRITICAL DIGITAL CITIZENSHIP**

RIBEIRO, Marlana Carla Peixoto. 2023. Discursos de ódio em comentários sobre postagens de celebridades femininas brasileiras no TikTok: Reflexões para uma Educação Linguística e Digital Crítica. 2023. 168 f. Dissertação (Mestrado em Educação, Linguagem e Tecnologias) – Universidade Estadual de Goiás, Anápolis, 2023.

ROJO, Roxane. (Org.). *Escola conectada: os multiletramentos e as TDICs*. São Paulo: Parábola, 2013.

RUEDIGER, Marco Aurélio; GRASSI, Amaro. (coord.). *Discurso de ódio em ambientes digitais: definições, especificidades e contexto da discriminação on-line no Brasil a partir do Twitter e do Facebook*. Policy paper, Rio de Janeiro: FGV DAPP, 2021.

SANTAELLA, Lucia. *Comunicação ubíqua: repercussões na cultura e na educação*. São Paulo: Paulus, 2013.

SANTAELLA, Lucia; CARDOSO, Tarcisio. Para entender a complexidade das redes. *Revista Flamecos*, Porto Alegre, v. 21, n. 2, p. 742-765, maio/ago. 2014. Disponível em: <https://bit.ly/3vnP18N>. Acesso em: 12 nov. 2021.

SARMENTO, Daniel. *A liberdade de expressão e o problema do Hate Speech*. In: D. Sarmento (Org.). *Livres e iguais: estudos de Direito Constitucional*. Rio de Janeiro: Lumen Juris, 2006.

SEIXAS, Rodrigo. A ecologia digital argumentativa: possibilidades e perspectivas para uma análise retórica da argumentação multimodal. *Revista da Abralin*, São Paulo, v. 20, n. 3, p. 918-937, 2021.

TRINDADE, Luiz Valério. Discurso de ódio nas redes sociais. In: RIBEIRO, D. (Org.). *Feminismos Plurais*. São Paulo: Jandaíra. 2022.

URZÊDA-FREITAS, Marco Túlio de. Sobre linguagens, corpos e (contra)políticas: praxiologias e letramentos queer em tempos sombrios. *Muitas Vozes*, Ponta Grossa, v. 9, n. 2, p. 706-736, 2020.

ⁱTerm used to refer to the move to an internet as a platform, attributed to Tim O'Reilly, in 2003.

ⁱⁱTerm that refers to those who, in some way, find themselves socially in a situation of dependence and/or disadvantage in relation to another, "majority" group; does not exactly refer to a number, but to a place of subalternity.

**HATE SPEECH IN COMMENTS ON SOCIAL MEDIA:
TOWARDS A CRITICAL DIGITAL CITIZENSHIP**

Corresponding author:

Marlana Carla Peixoto Ribeiro

Faculdade Evangélica de Goianásia

R. 18, S/N – Goianésia/GO, Brazil. Zip code 76380-000

marlanacpr@gmail.com

This is an open access article distributed under the terms of the Creative Commons license.

