Counterspeech humor for discursive justice

punctum.gr

BY: **Saleta de Salvador Agra**

ABSTRACT

Humor can be key in counteracting harmful discourse and addressing *conversational injustices*. This paper introduces the concept of counterspeech humor as the antithesis of *hate speech*, highlighting its ability to promote discursive justice. First, the central role of the counterspeaker is analyzed; second, three specific strategies of counterspeech humor are proposed: resignifying parody, bending irony, and intentionally blocking laughter. These forms of counterspeech humor are explored as effective responses to neutralize discursive harm and dismantle unjust practices. The effectiveness of counterspeech humor and its limitations in addressing harmful discourse are assessed through concrete examples. The paper emphasizes the reparative capacity of humor, showcasing its potential to offer constructive responses to hate. It concludes that counterspeech humor is a powerful tool for challenging injustices, transforming exclusionary contexts, and promoting social justice through creative use of language.

ARTICLE INFO:

Volume: **11**

Issue: **01**

Summer **2025**

ISSN: **2459-2943**

DOI: **10.18680/hss.2025.0006**

Pages: **115-133**

KEYWORDS:

Hate speech

Feminist humor

Counterspeaker

Semiotic laughter

1. Introduction: In the name of humor

In the name of humor, we perform and easily justify very different actions. I am also going to invoke humor here and take advantage of the performative versatility that characterizes it, to think about a type of humor that, rather than doing, dismantles, disarms, and deconstructs. A kind of humor that I will call *counterspeech*, since, paraphrasing and expanding on the famous title of John L. Austin's (1962) book, its objective

Punctum. International Journal of Semiotics | **11:**01**:**2025
ISSN 2459-2943 | DOI: 10.18680/hss.2025.0006 | **punctum.gr**

is to undo things with signs. More specifically, I will focus on undoing bad things through humorous signs, proposing *counterspeech humor* as an antonym of *hate speech* (Zollner 2024).

Although humor and hate represent different concepts, their antagonism is not always apparent. This is evident, for example, when we resort to humor to support and excuse hostile forms of speech. Thus, the common expression "It's just a joke" represents an attempt to conceal the hurtful meaning of specific comments, as if the mere intention could minimize the impact of the utterance (De Salvador 2023b). However, the connections between humor and hate are not limited to the use of these "get-out clauses" (Shahvisi 2023:53). Both notions share a common background, bringing them together in sociocultural and psychological terms and expanding their relationship beyond mere discursive justifications. This paper will show that for counterspeech humor to function adequately as a response to hate speech,[1] various contextual and social factors must influence its pragmatic felicity. That is, the impact and effectiveness of counterspeech humor depend not only on a specific linguistic structure or on the intention of the sender but also on deeply rooted power dynamics.

To develop this idea, I will begin by paying attention to the notion of *counterspeech humor*, resorting to the specialized literature for each of its components, and articulating both notions coherently and well-founded. Next, I will look at the counterspeaker. That is, the central role of the receiver who, on receiving the hate speech, can deactivate it. Without ignoring the effort such a task entails in epistemic and pragmatic terms, I will address three discursive strategies that function as humorous tools. Thus, I will focus on parody, irony, and a specific type of laughter, and I will take them as examples to reflect on possible cases of counterspeech humor.

These three instances will be presented from the feminist perspective. This choice is justified because these forms of humor stand out for their critical approach, proving the relevance of what is nowadays known as the "feminist philosophy of humor" (Marvin 2022:4). In fact, prominent feminist philosophers (Braidotti 1996; Butler 1990; Haraway 1985) have not only used and analyzed them as critical tools but also claimed their feminist application. Furthermore, feminist humor, whether parodic and ironic or not, can be understood as a counterspeech to the extent that it is constructed as discourse in opposition to certain hegemonic narratives. In this sense, feminist humor, as a form of *subversive humor* (Walker 1988; Gilbert 1997; Homes and Marra 2002; Kramer 2015), is an example of discursive resistance (De Salvador 2024) since, as I will explain, it works as a tool capable of counteracting sign structures of power. As Barba (2021:170) said, "[i]f one of the natural strategies of feminism is to invert meanings to

[1] Although the expression *hate speech* can encompass different uses, in this context I will understand it as the hostile use of linguistic expressions meant to do harm.

highlight the inequality they contain, feminist humor is undoubtedly the epitome of that gesture." In Crawford's (1997:129) words, "the political uses of humor by feminists indicate the potential of the humor mode to infiltrate and disrupt dominant." Therefore, we face a disruptive, counterspeech humor capable of generating what Paolucci and Caruana (2019:69) call *semiotic laughter*, that is, laughter capable of challenging the established social order. Finally, instead of focusing on the injustices caused and generated by humor, I will focus on its potential for undoing. Thus, this paper deals with a counterspeech humor that promotes discursive justice: a humor that has the power to respond to hate speech.

2. Counterspeech humor: Affinity network

A possible combination of humor and counterspeech reveals a harmony based on a common reflective structure, since both can function as potential second-order arguments. This means that both counterspeech and humor are built based on a previous speech. In the case of counterspeech, it is clear that its development depends on an initial argument or position that serves as the reference to be questioned. Following Eco (1984:8), something similar happens with humor for "[h]umor[2] is always, if not metalinguistic, metasemiotic: through verbal language or some other sign system it casts in doubt other cultural codes." This sequentially subordinated structure connects humor and counterspeech as possible responses to a previous message or code. From this perspective, the connection between them focuses on a shared metasemiotic structure.

However, not all humor is counterspeech, nor do all types of counterspeech resort to humor. Differences occur both in the message and the intention. Therefore, the purpose and content of the second-order argument will give rise to the classic dichotomy of humor: it either reinforces or subverts the preceding code. Thus, unlike counterspeech, which always arises as a response to hate speech, humor can reinforce and insist on the code that precedes it and reaffirm what is considered established in a sociocultural context. For example, as Bergson (1900) defends, humor can fulfil a corrective social function, strengthening existing social norms. This is where humor stops being subversive to embrace the status quo and preserve a specific sociocultural code. And also, in extreme cases, humor converges with hate speech if, through its actions, it enhances social conventions that discriminate and cause harm.

Far from this offensive and hurtful humor, what I will classify as counterspeech is precisely the type of humor that is presented to challenge and reverse situations

[2] In his paper, Eco (1984) draws a distinction between comedy and humor that we will not go into here, for want of space.

Punctum. International Journal of Semiotics | **11**:01:2025
ISSN 2459-2943 | DOI: 10.18680/hss.2025.0006 | **punctum.gr**

of *discursive injustice* (Kukla 2014). That is, counterspeech humor reacts to the unfair restriction that limits the linguistic capacity of certain speakers by preventing them from performing certain speech acts. In fact, counterspeech is always a way to combat linguistic toxicity (Tirrell 2018), which systematically harms some individuals just because they have a specific social identity. Undoing the damage of toxic speech is, therefore, the task of counterspeech. And it is when this is done through humor that we can see the counterspeech humor, whose main aim is precisely to deactivate toxic discourse. This does not imply understanding humor as a simple ingredient added to counterspeech and capable of facilitating or amplifying its effectiveness (although this may be true). In fact, humor can take the form of counterspeech.

Thus, in the following pages, rather than considering humor as a mere stylistic flourish, I will focus on a type of performative humor that defuses the harmful potential of the message that precedes it. With this purpose in mind, I will check my proposal of counterspeech humor against historical and contemporary theories of humor, seeking to identify a possible network of conceptual affinities. In other words, I will account for the possible similarities that, paraphrasing Wittgenstein (PI § 67), make up the "family resemblance" of humorous language games.

The meta-semiotic structure of humor focuses on a discourse capable of putting its own meaning in crisis. This undoing capacity of humor also depends on the purpose. If humor is counterspeech, its objective is to counter-argue to refute a previous thesis. A thesis that is necessarily circumscribed within the general framework of hate speech, understood as a type of speech that discriminates, causes harm, and systematically reifies oppressions. Therefore, counterspeech humor is presented as a type of response that seeks to deactivate the force of an oppressive discourse. A response that can be interpreted from the perspective of historical theories of humor (Morreall 1987), namely in terms of superiority, relief, or incongruity, or, in other words, considering the three elements activated by humor: authority, emotion, and the logic of discourse.

Thus, regarding the first theory, if we understand humor as a manifestation of superiority, counterspeech humor would seek to mock the interlocutor to disrupt their argument. In this case, humor could reverse the meaning of the previous argument – perhaps more aggressively and directly – and turn it into an object of laughter. There would be an attempt to ridicule the person and, consequently, to counteract their point of view, either by means of what we could call *ad hominem* humor or by mockingly discrediting the reasoning heard.

In light of the second theory, counterspeech humor could be analyzed as a form of relief, particularly when humor works as a release in uncomfortable and embarrassing situations, such as those generated by hate speech. Thus, we could get relief by laughing and, hence, reacting and dissipating the bad vibes of a tense and aggressive atmosphere.

Finally, from the perspective of the theory of incongruity, counterspeech humor manifests itself more obviously. This is so because this theory emphasizes the shock, the twist, and the challenge of previous expectations, bringing into question and testing our system of beliefs and habits. It is based on a contradiction between what is expected and what happens. From this point of view, the response of counterspeech humor may give way to an incongruity regarding the previous statement. However, incongruities can also generate a disruptive effect on the dominant discourse. In both cases, an attempt would be made to disrupt unfair presuppositions and highlight discursive inconsistencies while simultaneously introducing a new perspective into discourse. To sum up, making fun of the interlocutor, freeing oneself from an oppressive situation, or destabilizing expectations can be counterspeech humor resources, provided they help to dismantle discursive injustices.

The widely revisited and expanded three theories[3] have recently experienced what could be described as a political turn due to the assumption of a humor-*situated position*. This entails a critical perspective that explicitly addresses the political implications of humor, considering both the social context and the power relations that percolate through it. The focus is on how the experiences of the subjects involved and their position in a specific humorous situation influence their ability to make and receive humor. Among the current theories that align with this approach, the following conceptual proposals stand out: *charged humor* (Krefting 2014), *subversive humor* (Kramer 2015), or *humorwork* (Billingsley 2019).

All of them understand humor as a socio-political practice that "cannot be separated from the relationship between a subject and the political world in which they live" (Billingsley 2019:17). By politicizing humor, we also assume what the Willet and Willet (2019:18) call "humor from below." That is, a situational humor that seeks change and that "it would not just reinforce or temporarily invert hierarchies but level them" (Willet and Willet 2019:35). Thus, by using the adjective *subversive*, Kramer (2015:122) focuses on a type of humor that, as a persuasion device, seeks "to disclose and transmit information in such a way as to create change in both attitudes and practical social interactions through bringing to light flaws in our thinking and acting." The same occurs with the "charged" metaphor used by Krefting to describe a double humor movement, in which "[j]okesters unmask inequality by identifying the legal arrangements and cultural attitudes and beliefs contributing to their subordinated status – joking about it, challenging that which has become normalized and compulsory, and offering new solutions and strategies" (Krefting 2014: 2). Both "subversive" and "charged humor"

[3] The three historical theories have been reformulated several times. And new theories have also been proposed, such as the play theory and the theories focused on the linguistic-formal aspects of humor or its epistemic dimension.

Punctum. International Journal of Semiotics | **11:**01**:**2025
ISSN 2459-2943 | DOI: 10.18680/hss.2025.0006 | **punctum.gr**

make an emphasis on exposing the situation and proposing ways to transform it. Both would be "humorworks" that, far from idealized and essentialist positions,[4] are based on the embodied subject situated in the context of humor. In this regard, it could be said that we are facing proposals related to feminist *empathic humor*, which also seeks to level social hierarchies and inequalities from the standpoint of linguistic effectiveness.

Counterspeech humor is also the outcome of the political turn. This is why its "family resemblance" with the rest of the proposals is evident in its criticism of those theories that obscure the experiences of marginalized groups. Their kinship also sprouts from their explicit commitment to social change. In fact, the metasemiotic structure of counterspeech humor makes it possible to contradict a toxic discourse and, by opposing it, it also questions the conventions that support it, thus questioning the semiotic mechanisms whereby hate speech is constructed and exposing the underlying structure of power. Like the other proposals mentioned, counterspeech humor pays special attention to humor's socio-political and subjective structures, prioritizing them over the formal-linguistic features. Besides, it is in line with the growing philosophical relevance that counterspeech has recently acquired (Caponetto and Cepollaro 2022; Caponetto and Cepollaro 2024; Tirrell 2018; Langton 2018b; MacGowan 2018; Tirrell 2018; Ulmann and Tomalin 2024). It shares with counterspeech the urgency of offering a response that counteracts widespread hate speech. Counterspeech humor, as a tool for individual and collective use, is presented as an intentional act that seeks to ruin the "felicity" of the preceding speech act. That is to say, it aims to prevent what has been told from having an impact; in other words, it tries to deactivate its harmful effects. So, the "felicity conditions" of the speech act announced by Austin (1975:14-15)[5] remain in the hands of the receiver (De Salvador 2024), that is, of the subject who seeks to contradict through humor.

To clarify the concept of counterspeech humor, I will begin by exploring the role of the counterspeaker (2.1), as I consider their position pivotal in generating the humorous effect. I will then analyze three strategies to produce such an effect (3).

---

[4] Thus, Billingsley (2019:47) writes:

> core essentialist approaches to humor such as those found in Victor Raskin's script semantic theory and Raskin's and Salvatore Attardo's general theory of verbal humor has staved off a consideration of humor's relationship to power, prejudice, social norms, and subjectivity, instead calling for an investigation of idealized subjects who navigate a purified linguistic network shorn from any power-laden social context.

[5] The philosopher mentioned four key factors, grouped into three types of *felicity* (A, B and Γ), for the success of broadcasts that take into account what has been said, how it has been said, who has said it and the context of enunciation.

2.1. Counterspeaker humor: Agency and authority

The figure of the counterspeaker is that of an active receiver who, like the Barthesian (1977) reader, plays a central role in the construction of meaning. His/Her participation in the speech context positions him/her not only as a co-creator of meaning but also as someone with the capacity to question and redirect conversation. From this second perspective, we could say that the counterspeaker assumes the role of Eco's (1986:142) semiological guerrilla for s/he must "discuss the arriving message in the light of the codes at the destination, comparing them with the codes at the source." Code interplay (the confrontation between the entry and exit codes) allows for generating normative changes within a given speech situation. Such changes would allow the rules of the conversational game to be altered, undermining the direction of discourse and the position of whoever assumes authority in it. In this scenario, humor, when it is *guerrilla humor*, is also situated in an *intentio receptoris* capable of opening spaces for discursive resistance in the hands of the recipient. Thus, the counterspeaker uses humor to deliberately confront hate speech. Its comic action, far from focusing on entertainment, is aimed at disarming and re-signifying, generating new meanings that question unfair presuppositions and challenge oppressive ideologies. To achieve this goal, the counterspeaker often uses the indirect mode of humor to shock and promote change. As Crawford (1997:134) says, "Indirect modes can save face, minimize accountability for one's actions, and slip taboo topics into conversation. And humor is perhaps the most flexible and powerful of indirect modes." Humor thus emerges as a promising counterspeech weapon since, in addition to protecting the comedian, it captures the audience's attention and makes it easier to join the conversation. These ingredients, specifically intended to respond to hate speech, transform humor into a potential tool for semiological guerrilla warfare.

If hate speech is an invitation to leave the conversation (Langton 2018b), its antithesis, here, counterspeech humor, can become a way to join it. Unlike other forms of speech, humor allows marginalized groups to access discourse more easily and even claim spaces of agency. This is so because the indirect mode is a mitigated way of softening the impact of the message, making it more socially acceptable, so the humor

> can be a useful strategy in interpersonal conversation, where a humorous mode of communication may give a speaker a chance to demonstrate her point of view more vividly, and with less cognitive resistance on the part of her interlocutors, than serious assertion. (Butterfield 2022:700)

For speakers whose linguistic agency is systematically restricted, humor can become a key tool to access one's conversational exchange. An example of this is found in the first-person account of the Australian comedian Hannah Gadsby, who, facing

Punctum. International Journal of Semiotics | **11:**01**:**2025
ISSN 2459-2943 | DOI: 10.18680/hss.2025.0006 | **punctum.gr**

prejudice for being lesbian and fat and for having Autistic Spectrum Disorder, describes how, from her childhood and on numerous occasions, humor became the way to access discourse. Thus, for example, Gadsby (2022:107) explains that she would try to join any conversation "the only way I knew how: by saying something funny." This does not mean that humor, as a form of linguistic agency vindication, is a tool that is easy to use, unambiguous, or free of epistemic efforts.[6] Nor does it mean it is trouble-free or would never backfire and reinforce oppressions. In this sense, the examples we found show both the inherent instability of intentions (incapable of unilaterally determining a speech act) and the prejudices associated with the subject of humor itself. A subject constructed in light of a stereotype that defines both the image of the humor producer and that of its object. Thus, traditionally, men have occupied the leading role, while the secondary role has been assigned to women (Bunkers 1985; Bergmann 1986; Walker 1988).

Joining the conversation by using humor can be risky if you are not an authorized subject, since you could suffer what Butterfield (2022), based on Fricker's (2007) epistemic injustices, calls "comedic hermeneutical injustice." Such an injustice systematically "manifests itself in disadvantages, experienced by members of hermeneutically marginalized groups, in successfully sharing humor and being considered funny by others" (Butterfield 2022:688). The ability of hate speech to undermine the illocutionary force of the words uttered by specific individuals turns re-entering through humor into a difficult task. In fact, the subject targeted by hate sometimes resorts to self-mockery or self-ridicule (i.e., making humor at one's own expense) as a strategy to enter or reclaim the conversation. We are faced with a form of self-deprecation characteristic of *marginal humor* (Gilbert 2004) that plays with stereotypes[7] and blurs the fine line between laughing at oneself and laughing with someone. This is precisely one of the reasons why Gadsby, in her acclaimed stand-up comedy, *Nanette* (2018), announces that she is leaving comedy. In this comic monologue, the comedian openly questions the rule of self-hatred imposed on marginalized subjects as a way of accessing discourse and, by extension, comic authority. This would be a case of "charged humor," where the complaint of becoming a joke is intertwined with criticism because it "reveals one's immediate experience of second-class citizenship and gives us proactive means of addressing inequality" (Krefting 2014: 26). Authority in discourse, however, can also be

[6] It not only requires creativity and wit, cognitive skills but also the ability to quickly apply a range of semantic and pragmatic abilities.

[7] Regarding the transgressive role that stereotypes can play in stand-up comedy humor, I refer to the study by Gilbert (2004:165), who stated:

> By performing marginality onstage, a comic is doubly empowered -first by foregrounding difference, and second by commodifying and ultimately profiting from that difference. In this way, social stigma may function as a rhetorical means to a political end.

achieved through other means. This is precisely what constitutes the core of counterspeech, since it implies breaking with the authority assumed by a subject who carries out a linguistic act of hate. Thus, the counterspeaker seeks to provoke a normative change that counteracts the power of the said subject, challenging both their authority and the content transmitted.

As we will see below, the different strategies of counterspeech focus on undoing unfair discursive practices, deactivating the message, and, therefore, disavowing it. Counterspeech in the hands of the comedian reflects the intention not only to participate in a conversation but also to exercise an agency that responds to a situation of linguistic oppression. Besides, the situation becomes even more counterdiscursive when those seeking to join are subjects in socially subaltern positions. Hence, counterspeech humor can be read as a vindication of comic agency and a challenge against the discursive authority that attempts to expel them.

3. Some strategies of counterspeech humor

The different counterspeech strategies depend on several factors, such as the context of the communicative situation, the lexicon used, the content of the counterargument, or the social position of the interlocutors (Caponetto and Cepollaro 2024). Similarly, the effectiveness of counterspeech humor is also conditioned by the message to which it responds and by the place and people involved in the conversation. With regards to the latter, the fact that the counterspeaker is a member of the hate speech target group is highly relevant because, aside from the aforementioned epistemic injustices – as MacGowan (2018) denounces – on the one hand, "it places the burden of the remedy on those targeted (and potentially harmed) by the allegedly harmful speech" (MacGowan 2018:183) and, on the other hand, "it appears to assume a level playing field for all speakers" (MacGowan 2018:185).

In addition to the position of the counterspeaker, the context also exerts an influence. A stand-up comedy-style monologue is not the same as an everyday conversation. In the case of the monologue, humorous authority is influenced by market logic, since, as Krefting (2014) explains, some comic characters are more marketable than others. This limits activist humor, such as charged humor, by confining it to a narrower market niche. However, the stage continues to be a space for enunciation that grants the comedian some power: a place to express oneself and contradict hate speech. This is precisely how Gadsby (2022:251) describes her position after announcing, on stage, that she is leaving comedy: "My stunned audience remained so still and quiet, I had them in the palm of my hand, and that's when I first really understood that despite having little power in the world, I did have incredible power over a room when I was on stage. The time had come to wield it." The situation off stage changes, and with it,

Punctum. International Journal of Semiotics | **11:**01:2025
ISSN 2459-2943 | DOI: 10.18680/hss.2025.0006 | **punctum.gr**

so do the dynamics of authority. In fact, in everyday conversation, those who cause verbal harm usually rely on authority,[8] whether they have it or not, to spread and legitimize prejudices and stereotypes. Therefore, as in guerrilla warfare, counterspeech humor tactics must focus on delegitimizing that authority.

Disavowing a hate message without resorting to legal and institutional censorship can be achieved in different ways. Each of them constitutes a different type of counterspeech. However, they all share the objective of undoing a verbal expression that creates hostility towards people because they belong to a discriminated social group. In this context, counterspeech humor arises with the same purpose: addressing situations of patent injustice by undermining the illocutionary force of a previous message, that is, by ruining the effectiveness of the harmful speech act. Among the ways in which humor deploys its counterspeech potential, I will first take a brief look at parody as a way to imitate and transform meanings; then, I will look at irony as a resource that exploits double meaning and implicit contradiction; and, finally, I will focus on laughter as a sound that is capable of interrupting and preventing the percolation of unfair presuppositions. The approach is that of a feminist humor whose political commitment shows strategically since "feminism has shown that humor can also be the perfect Trojan horse to enter a society besieged by structural male chauvinism" (Barba 2021:14). It is about entering and ruining the sexist, racist, xenophobic, classist, homophobic party and any other party celebrating discrimination and perpetuating social inequalities. In other words, it is about practicing what would be typical of "killjoy" humor (to make a pun based on Ahmed's 2023 terms). Or, using Clinton's terminology, fusing *humor* and *feminist*, the *fumerist*, because, in Barreca's (2023:178) words, this term "captures the idea of being funny and wanting to burn the house down all at once." For "feminism as fumerism offers one way to confront and detoxify the stereotypes, to joyfully reappropriate energy and eros from systems of domination" (Willet and Willet 2019:45). In short, using humor to bring about resignification, bending and blocking, or, in other words, using parody, irony, and laughter to stop hate speech.

3.1. **Parodying to achieve resignification**

Humorous counterspeech games exploit the idea that several meanings coexist in a single signifier. As a form of humor, parody plays with that distance between the elements of the sign, stretching it until it breaks and thus undoing the original union. This is where the parodic response exhibits its counterspeech potential. That is, by parodying, it is possible to alter the conversation, introduce new information and even undo the relationship between the signifier and its harmful semantic load. This humorous strategy, based on imitating and exaggerating until meaning is distorted, was adopted and promoted

[8] As to authority in hate speech, I refer to Langton (2018a).

by the feminist movement, especially in the early nineties of the last century. Thus, at a theoretical level, authors such as Butler (1990) or Braidotti (1996) highlighted the political potential of parody, while artistic collectives, such as the Guerrilla Girls or the broad phenomenon of the Grrls, and cyberfeminist groups, such as the VeNuS Matrix, put it into practice. The common idea was, on the one hand, to reveal the instability of a dominant semiotic configuration and, on the other hand, to combat the hate speech that was transmitted inside and outside cyberspace, ridiculing it to the extreme.

Butler (1997:15) maintained that the relationship between particular words or speeches and their offensive charge is neither fixed nor invariable. In fact, she thought of language as an iterable activity open to change where repetition opened a possible gap between the harmful speech act and its perlocutionary effects. A repetition that Bergson (1900) had previously highlighted as a comic element with the capacity to generate a break with whatever is expected, and that Butler, in line with Derrida, picks up to claim a linguistic agency capable of re-signifying.: "[t]he resignification of speech requires opening new contexts, speaking in ways that have never yet been legitimated, and hence producing legitimation in new and future forms" (Butler 1997: 41). From this point of view, the parodic counterspeaker can resort to generalizations, stereotypes or even prejudices and use them as weapons with which to dismantle their simplifying potential. Drag culture is one of the examples that Butler (1990) gives as a type of gender exaggeration that dissolves associations through parodic hyperbole. We find a similar case in the nascent cyberspace when the VeNuS Matrix group created, in 1993, the interactive game *All New Gen*. This gamification takes up established meanings about bodies, acting as an iconic and cartoonish counterspeech regarding gender norms. Thus, they present avatars with hypersexualized bodies that are dismantled to challenge biological constructions (De Salvador 2018).

Parody and acting "as if" intersect when imitation has a critical purpose. This is what Braidotti (2002:353) suggested when he stated that "parody can be politically empowering on the condition of being sustained by a critical consciousness that aims at the subversion of dominant codes". The philosopher identified the politics of parody with the transgressive aesthetics of the Guerrilla Girls, with their King Kong masks, and with the roar of the Riot Grrl (an onomatopoeia that imitates the wild sound by changing the *i* for the *r*). Their counterspeech humor can be clearly seen in the poster that the Guerrilla Girls put up in the Metropolitan Museum in 1989, where they openly parodied the painting *La Grande Odalisque* (1814). The naked woman, dressed in a gorilla mask, was accompanied by a written text that read: "Do women have to be naked to get into the Met. Museum? Less than 5% of the artists in the Modern Art Sections are women, but 85% of the nudes are female." The parody of the Guerrillas sought to call the iconic significance of the naked woman into crisis, though reappropriated to transgress gender codes semiotically. In contrast to the passive and sexualized figure that

Punctum. International Journal of Semiotics | **11:**01:2025
ISSN 2459-2943 | DOI: 10.18680/hss.2025.0006 | **punctum.gr**

Ingres presents for the male gaze, the Guerrilla Girls' parody exposes and critiques the mechanisms of silencing (De Salvador 2023a), invisibility, and exclusion that women have historically faced within the artistic canon. Thus, by deliberately exaggerating, the strategic procedure of parody can become a direct counterspeech mockery, making it a more dangerous situation for the counterspeaker.

3.2. **Using irony for bending**

Humorous shortcuts are essential to avoid direct confrontation, especially in a situation of verbal violence. The Guerrilla Girls masks focused on the message they wanted to convey, but, at the same time, they turned out to be a good protective shield for them to hide behind. An anonymity that cyberfeminists themselves also exploited to their advantage in cyberspace and that, to this day, continues to be valid, for example, in humorous formats such as memes. Tools that, in any case, in the hands of the counterspeaker must be used to counteract narratives based on inequalities and expulsions precisely. When it is counterspeech, parody re-signifies toxic discourse, using the possibilities that acting "as if" opens. In fact, what Braidotti (2002:353) calls "the philosophy as if", in his own words, "with its ritualized repetitions, needs to be grounded in order to be politically effective." The "as if" practices can take on a parodic but also ironic character, or a simultaneous combination of both. In this sense, the clearest example is found in *A Cyberfeminist Manifesto for the 21st Century* (1991) and *Bitch Mutant Manifesto* (1996), where the VeNuS Matrix combined satire, parody, and irony in equal parts. They present themselves with phrases such as "We are the virus of the new world disorder" (VeNuS 1991) and "We are the malignant accident which fell into your system while you were sleeping" (VeNuS 1996), which exemplify this mixture and through which they denounce the discursive injustice inherent in the patriarchal and technological system. In the two *Manifestos*, they follow Haraway's approaches to the cyborg as fiction, while openly adopting her conception of irony. As Haraway (1991a:149) points out, "[i]rony is about humor and serious play. It is also a rhetorical and political strategy, one I would like to see more honored within socialist-feminism. At the center of my ironic faith, my blasphemy, is the image of the cyborg." Later, she wrote, "[f]or many of us, the irony made it possible to participate-indeed, to participate as fully committed, if semiotically unruly" (Haraway 1991b:318). Using irony as a critical access tool manifests itself in counterspeech through tactics such as *bending* or *disunderstanding*.

Caponetto and Cepollaro (2022) proposed their own modality, which aligns with current studies on counterspeech, in which toxic speech is counteracted through a deviant and improved response. A response that deviates from the path demarcated by the original discourse, twisting it until its toxicity is transformed or, in other words, "turning poison into medicine." Thus, the authors above point out that

> [q]ua distorting move, [bending] succeeds when, in addition, it manages to make the ameliorated content enter the common ground in place of the toxic content. This is the case when the toxic speaker plays along and does not retort by explicitly asserting the toxic content they were implicitly conveying. (Caponetto and Cepollaro 2024:71)

Bending, using Tirrell's (2018) terminology, would be both an inoculation that reorients what has been said to generate a positive effect – neutralizing the toxic damage previously inflicted in a conversation – and a kind of vaccine that prevents the installation of toxins in the "common ground," by stimulating the immune system and stopping the spread of discursive toxins. In other words, the "distortion in the service of social justice" (Caponetto and Cepollaro 2022:3). In this sense, bending would also be an example of what Maxime Lepoutre (2019:34) calls "positive counterspeech" since it "is less about directly contesting a distorted vision of the world, and more about affirming a correct vision of the world that is inconsistent with the falsehoods at hand."

Diverting the course of a conversation can also be carried out by *disunderstanding*. The pun on words created by Shahvisi (2023) points to a deliberate misunderstanding with which the listener refuses to accept an explanation. With this discursive movement, the person is forced to openly expose the assumptions that underlie a more than dubious explanation. The reason for this is that

> [e]xplanations for marginalizing actions or perspectives often rely on vague, euphemistic phrases. When a person is forced to explicitly state their assumptions, they must reflect more closely on what they're actually saying, and whether or not they can justify it. (Shahvisi 2023:136)

This way, the counterspeaker, with a deep pedagogical intention in the style of Socratic irony, feigns ignorance to make the interlocutor fall into an open explanation of whatever was implicitly said. This is an ironic dissimulation, a maneuver to put unfair presuppositions on the table, thus opening the possibility of confronting or blocking them. Therefore, disunderstanding begins with an ironic pretense, as in the case of bending, to finally separate from it and seek an open disagreement, as occurs with the last counterpeech strategies I will expose next: *blocking*.

3.3. Laughing for the sake of blocking

Jankélévitch argued that irony, when used with humor, can make deep philosophical tensions more friendly and accessible. In this sense, humorous irony does not lose its critical capacity during bending or disunderstanding, because it is open. That is to say, "humor is open irony: because while closed irony is not aimed at instructing the other,

Punctum. International Journal of Semiotics | **11:01:2025**
ISSN 2459-2943 | DOI: 10.18680/hss.2025.0006 | **punctum.gr**

open irony is ultimately the beginning of understanding and spiritual community" (Jankélévitch 1964:172). Irony, at the service of humor, does not aim to devalue the other, but rather to promote understanding as a basis for building community. In contrast to historical theories of superiority, humorous irony embraces what Willet and Willet (2019:10) renamed *leveling humor*: an empathetic humor from the margins "to reject the old patriarchal formulas for laughs" (Willet and Willet 2019:151). In other words, compared to the Hobbesian (1969) laughter typical of the theory of superiority, a horizontal laughter is proposed, a laughter that, instead of humiliating and reinforcing inequalities, promotes new power dynamics based on complicity between people.

Laughter can arise as a reaction to a comical situation but can also originate in opposite circumstances. It is a complex response that can be read both from what causes it and from what it generates. Or, in Austin's terms, there is perlocutionary laughter as an effect or cause, and illocutionary laughter that does things with sounds. A laughter, which I call *counterspeech*, emerges from the broad framework of the illocutionary force of laughter. A laughter that could act as a blocking (Langton 2018b) and that, with its sound, causes speech to stop. Butler (2019) highlighted the political potential of laughter as a noise capable of disrupting discursive functioning because "laughter interrupts discourse" and "exercises a renewing power." In this sense, laughter intervenes in communication as a side effect and an active response to what is heard. So, as Heller (2005:25) points out, "[l]aughter is judgmental; with it, we pass a judgment." When we laugh, we make judgments, assess, and even challenge what we hear. In this context, a counterspeech laugh deliberately prevents unfair presuppositions from being inserted into the discourse, paralyzing violence and functioning as a strategy similar to blocking, as Langton (2018b) proposed.

In Langton's proposal (2018b), blocking is presented as a semiotic-political tactic aimed at explicitly preventing the "accommodation of injustice." It is a counterspeech tool that combines Lewis's (1979) *rule of accommodation* with Sbisà's (2021) idea of *explicitation practices.* So, on the one hand, "[w]hen you block something, you don't 'accommodate' it – you don't adjust to it, or help it along" (Langton 2018b:145) and, on the other hand, "[w]hen the hearer blocks with explicitation, forcing the speaker's cards onto the table, the back-door speech acts fail, at least as *back-door* speech acts" (Langton 2018b:154). In both cases, the focus is on the presuppositions, that is, on the implicit contents that are communicated and, more specifically, on those that perform a persuasive function. Thus, it is about "assumptions, not necessarily conscious but liable to be brought to consciousness, about how our human world is and how it should be" (Sbisà 1999:493). Blocking implies preventing implicit content from percolating through the speaker's intervention, thus rejecting ideological presuppositions that transmit derogatory and pejorative content. To carry out the act of blocking, the presupposition must be made explicit, since, as Sbisà (1999:507) states,

> [o]nce it is made explicit, of course, a presupposition becomes just an assertion and, like all assertions, whilst the speaker is committed to defending it, the listener or reader is allowed to challenge it and even to ask the speaker for his or her grounds for making that assertion.

Langton (2018b:149) points to other forms of blocking in addition to explicitation since "[s]ometimes a rephrasing, a raised eyebrow, or a joke, might work better than a 'Wait a minute!' or a righteous calling out." Laughter could also act as an additional form of non-explicit blocking – one where it becomes an obstacle in the speech of someone who seeks to hurt through their words. Although laughter is not an explicit form of articulate verbal communication, it can convey many meanings. Words are unnecessary, as they can function as an acoustic response that shows complicity or disagreement. When the expected perlocutionary effect is for someone to laugh, not doing so does not nullify the act but fails as it provokes an unintended reaction. Now, laughing also can illocutionally annul an act, and that is why it could be interpreted as a powerful tool to stop and block a message. In this second sense, all forms of killjoy laughter may serve as illustrative examples of discursive blocking. As Ahmed (2023:174) writes, "laughter can be the sound of a door being slammed," thus preventing anything from slipping through it, as Langton (2018b) suggests. Although historical records of women laughing are scarce – a fact that is revealing in itself – the well-known biblical figure Sarah stands out as an act of resistance. Her laughter not only expresses skepticism but also disrupts the underlying presuppositions of the discourse directed at her: that her body is still fertile, that motherhood remains a possible destiny despite her advanced age, and that she is expected to accept the divine promise without question. Also, laughter showing teeth –which denotes both aggression and disobedience – can break the conventional structure of the conversation, introduce changes in the assumed normative positions, and, thereby, stop toxic discourse. In this sense, the obscene and transgressive laughter of Baubo, a figure from Greek mythology, together with her irreverent gesture, constitutes another paradigmatic example: it manages to break Demeter's silence and reset the discursive order, thus subverting the imposed social expectations. These are forms of laughter not covered by the hand, blocking laughter that, like ironic or parodic laughter, acts as counterspeech, challenging and embarrassing speakers who seek to harm with their speech. In short, semiotic laughter emerges, and this, as Paolucci and Caruana (2019) taught us, can call into question the social system and, therefore, could contribute to discursive justice.

Punctum. International Journal of Semiotics | **11:**01:2025
ISSN 2459-2943 | DOI: 10.18680/hss.2025.0006 | **punctum.gr**

4. Conclusion: Semiotic laughter for discursive justice

Neither parody, nor irony, nor laughter is *per se* a counterspeech movement. In fact, they could all be humorless and hateful. This shows humor's political instability (Billingsley 2019) and counterspeech (MacGowan 2018), since they are not an infallible formula. They are sign practices, and their use, as Wittgenstein (1958) said, determines their meaning. Now, this approach opens up the possibility of assigning humor the ability to bring about resignification, bending, or even blocking oppressive meanings. In doing so, humor becomes a counterspeech medium, capable of altering dominant discourses and generating new interpretations. By intervening in the semiotic process, humor challenges social conventions and offers ways to question and transform established power structures that generate inequalities.

As has already been said, responding with humor is not an infallible remedy, nor is it a strategy that guarantees counteracting the verbal damage caused. However, as they say, sometimes, laughter is the only way out. In fact, when laughter is semiotic and counterspeech, it can have a contagious illocutionary force that may defuse tensions, resolve conflicts, and even offer respite in an environment of pressure and oppression. It is then that laughter is presented as a form of discursive resistance that, just like the parodic and ironic tactic, proves its ability to undo.

However, counterspeech humor lives during a conversation, so, despite seeking to stop hatred, it could be followed by a rejoinder. In fact, one of the risks of responding with humor is that it can lead to everything becoming trivial and losing seriousness in the face of unfair situations. This is so because, as was mentioned at the beginning of this paper, many things can be done in the name of humor. Here I have sought to accentuate the good things that can be done with it, those that contribute to a discursive justice that, instead of silencing and oppressing, favors the entry of other voices, especially those that have seen their illocutionary potential undermined.

References

Ahmed, Sara 2023. *The Feminist Killjoy Handbook*. London: Penguin Random House.

Austin, John L. 1975 [1962]. *How to Do Things with Words*. Cambridge: Harvard University Press.

Barba, Andrés 2021. *La risa caníbal. Humor, pensamiento crítico y poder*. Barcelona: Alpha Decay.

Barreca, Regina 2013 [1991]. *They Used to Call Me Snow White… But I Drifted: Women's Strategic Use of Humor.* Hanover: University Press of New England.

Barthes, Roland 1977. The death of the author. In: *Image-Music-Text*. London: Fontana Press, 142-148.

Bergson, Henry 1913 [1900]. *Laughter. An Essay on the Meaning of the Comic*. New York: The Macmillan Company.

Bergmann, Merrie 1986. How many feminists does it take a joke? Sexist humor and what's wrong with it. *Hypatia* 1 (1): 63-82.

Braidotti, Rosi 2002 [1996]. Cyberfeminism with a difference. In: Mui, Constance L. and Murphy, Julien S. (eds.) *Gender Struggles: Practical Approaches to Contemporary Feminism*. Lanham: Rowman and Littlefield, 347–357.

Billingsley, Amy 2019. *Humorwork, feminist philosophy, and unstable politics*. Ph.D. thesis, Department of Philosophy, University of Obregon.

Bunkers, Suzanne 1985. Why are these women laughing? The power and politics of women's humor. *Studies in American Humor* 2 (4): 82-93.

Butler, Judith 1990. *Gender Trouble. Feminism and Subversion of Identity*. New York: Routledge.

Butler, Judith 1997. *Excitable Speech. A Politics of the Performative*. New York: Routledge.

Butler, Judith 2019. *Out of Breath: Laughing, Crying at the Body's Limit*. https://vimeo.com/352083590

Butterfield, Paul 2022. Comedic hermeneutical injustice. *Hypatia* 37 (4): 688-704.

Caponetto, Laura and Bianca Capollaro 2022. Bending as counterspeech. *Ethical Theory and Moral Practice* 26 (4): 577-593.

Caponetto, Laura and Bianca Capollaro 2024. The philosophy of counter-language. In: Ullmann, Stefanie and Tomalin, Marcus (eds.) *Counterspeech. Multidisciplinary Perspectives on Countering Dangerous Speech*. New York: Routledge, 56-71.

Crawford, Mary 1997. On conversational humor. In: *Talking Difference. On Gender and Language*. London: Sage, 129-170.

De Salvador Agra, Saleta 2018. Guerrillas ciberfeministas. *Revista Andaluza de Antropología* 14: 133-153.

De Salvador Agra, Saleta 2023a. Cómo no hacer cosas con palabras y cómo hacerlas con el silencio. *Endoxa: Series Filosóficas*.52: 143-162

De Salvador Agra, Saleta 2023b. 'Mujer, era una broma.' Deshacer cosas con palabras. *Revista Ártemis* 35 (1): 10-25.

De Salvador Agra, Saleta 2024. Resistencias discursivas: Resignificar, bloquear y desviar. *Logos: Revista de Lingüística, Filosofía y Literatura* 34 (2): 742-762.

Eco, Umberto 1986 [1967]. Towards a semiological guerrilla warfare. In: *Travels in Hyper Reality*. New York: Harcourt, Brace, Jovanovich, 135-145.

Eco, Umberto 1984. The frames of comic 'freedom'. In: Marcia E. Erickson (ed.) *Carnival!* Berlin: De Gruyter, 1-11.

Fricker, Miranda 2007 *Epistemic Injustice. Power and the Ethics of Knowing*. Oxford: Oxford University Press.

Gadsby, Hannah 2018. *Nanette*. Netflix.

Punctum. International Journal of Semiotics | **11:**01:2025
ISSN 2459-2943 | DOI: 10.18680/hss.2025.0006 | **punctum.gr**

Gadsby, Hannah 2022. *Ten Steps to Nanette. A Memoir Situation*. New York: Penguin Random House.

Gilbert, Joanne R. 1997. Performing marginality: Comedy, identity, and cultural critique. *Text and Performance Quarterly* 17: 317-330.

Gilbert, Joanne R. 2004. *Performing Marginality. Humor, Gender, and Cultural Critique.* Detroit, MI: Wayne State University Press

Haraway, Donna J. 1991a [1985]. A cyborg manifesto: Science, technology, and socialist-feminism in the late twentieth century. In: *Simians, Cyborgs, and Women. The Reinvention of Nature*. New York: Routledge, 149-181.

Haraway, Donna J. 1991b. The promises of monsters. In: Lawrence Grossberg, Cary Nelson and Paula Treichler (eds.) *Cultural Studies*. New York: Routledge, 295-337.

Heller, Ágnes 2005. *Inmortal Comedy. The Comic Phenomenon in Art, Literature, and Life.* Oxford: Lexington Books.

Hobbes, Thomas. 1969. *The Elements of Law Natural and Politic*. London: Frank Cass.

Holmes, Janet and Meredith Marra 2002. Over the edge? Subversive humor between colleagues and friends. *Humor. International Journal of Humor Research* 15 (1): 65-87

Jankélévitch, Vladimir 1964 [1936]. *L'ironie*. Paris: Flammarion.

Kramer, Chris A. 2015. *Subversive Humor.* Dissertation: Marquette University. DOI: 10.13140/2.1.2061.1368

Krefting, Rebecca 2014. *All Joking Aside. American Humor and Its Discontents*. Baltimore: John Hopkins University Press.

Kukla, R. Quill 2014. Performative force, convention, and discursive justice. *Hypatia* 29 (2): 440-457.

Langton, Rae 2018a. The authority of hate speech. In: John Gardner and Leslie Green (eds.) *Oxford Studies in Philosophy of Law*. Oxford: Oxford University Press, 123–152.

Langton, Rae 2018b. Blocking as counter-speech. In: Daniel Fogal, Daniel W. Harris and Mat Moss (eds.) *New Work on Speech Acts*. Oxford: Oxford University Press: 144-164.

Lewis, David 1979. Scorekeeping in a language game. *Journal of Philosophy Logic* 8: 339-359.

Lepoutre, Maxime 2019. Can 'more speech' counter ignorant speech? *Journal of Ethics and Social Philosophy* 16 (3): 155-191.

MacGowan, Mary Kate 2018. Responding to harmful speech. In: Rebecca C. Johnson (ed.) *Voicing Dissent*. Abingdon: Routledge, 182-200.

Marvin, Amy 2022. Feminist philosophy of humor. *Philosophy Compass* 17 (7): 1-10.

Morreall, John 1987. *Philosophy of Laughter and Humor*. New York: State University of New York Press.

Paolucci, Claudio and Fausto Caruana 2019. Riso e logos. Il riso semiotico come proto-linguaggio, tra emozioni e socialità. *Rivista Italiana di Filosofia del Linguaggio* 66-77.

Shahvisi, Arianne 2023. *Arguing for a Better World. How Philosophy Can Help Us Fight for Social Justice*. New York: Penguin Books.

Sbisà, Marina 1999. Ideology and the persuasive use of presupposition. In: Jef Verschueren (ed.) *Language and Ideology. Selected Papers from the 6st International Pragmatics Conference*. Antwerp: International Pragmatics Association, 492-509.

Sbisà, Marina 2021. Presupposition and implicature: Varieties of implicit meaning in explicitations practices. *Journal of Pragmatics* 182: 176-188.

Tirrell, Lynne 2018. Toxic speech: Inoculations and antidotes. *The Southern Journal of Philosophy* 56 (1): 116-144.

Ullmann, Stefanie and Tomalin, Marcus (eds.) 2024. *Counterspeech. Multidisciplinary Perspectives on Countering Dangerous Speech*. New York: Routledge

VeNuS Matrix 1991. *A Cyber Feminist Manifesto for the 21st Century*. https://vnsmatrix.net/projects/the-cyberfeminist-manifesto-for-the-21st-century

VeNuS Matrix 1996. *Bitch Mutant Manifesto*. https://vnsmatrix.net/projects/bitch-mutant-manifesto.

Walker, Nancy A. 1988. *A Very Serious Thing. Women's Humor and American Culture.* Minneapolis: University of Minnesota Press.

Willet, Cynthia and Julie Willet 2019. *Uproarious. How Feminist and Others Subversive Comics Speak Truth*. Minneapolis: University of Minnesota Press

Wittgenstein, Ludwig 1986 [1953]. *Philosophical Investigations*. Oxford: Basil Backwell.

Zollner, Sebastian 2024. Counterspeech practices in digital discourse: An interactional approach. In: Stefanie Ullmann and Marcus Tomalin 2024 (eds.) *Counterspeech. Multidisciplinary Perspectives on Countering Dangerous Speech*. New York: Routledge, 27-56.

AUTHOR

Saleta de Salvador Agra is Associate Professor of Logic and Philosophy of Science, University Complutense of Madrid, Spain.