Exactly, what do you mean?*

Exactamente, ¿qué quieres decir?

*MANUEL ALMAGRO***
*NEFTALÍ VILLANUEVA****, *****

Abstract: The purpose of this paper is to explore the boundaries of a subset of the evaluative uses of language: offensive speech. More in particular, our goal is twofold. Firstly, we want to chart the relationship between context and evaluative utterances, as it stands in the current literature. Secondly, we focus on the experimental study of a particular interaction between contextual information and our evaluative claims –when the context is able to turn a seemingly descriptive utterance into an evaluative one. For this second purpose, we argue, certain recent positive proposals, in spite of their merit, come a bit short.

Keywords: Evaluative language, Offensive speech, Affective attitudes, Context-dependence, Expressive content.

Resumen: El propósito de este trabajo es explorar los límites de un subconjunto de los usos evaluativos del lenguaje: el discurso ofensivo. Nuestro objetivo es doble. Primero, introducimos la relación que hay entre el contexto y las proferencias evaluativas, tal y como puede rastrearse en la literatura reciente acerca de la cuestión. Segundo, nos centramos en el estudio experimental de una interacción particular entre la información contextual y nuestras afirmaciones evaluativas: cuándo el contexto es capaz de convertir una proferencia aparentemente descriptiva en una evaluativa. Para este segundo propósito, argumentamos, ciertas propuestas positivas recientes, a pesar de su mérito, son insuficientes.

Palabras Clave: Lenguaje evaluativo, Discurso ofensivo, Actitudes afectivas, Contexto-dependencia, Contenido expresivo

** FPI Predoctoral Researcher, Department of Philosophy I, University of Granada (Spain). Email: malmagro@ugr.es. My work focuses on political polarization, mental state attribution and the evaluative use of language. Recent publications: Almagro, M., Navarro, L. & Pinedo, M. (forthcoming), «Is testimonial injustice epistemic? Let me count the ways», *Hypatia: A Journal of Feminist Philosophy*, and Almagro, N. & Villanueva, N. (2021), «Polarización y tecnologías de la información: radicales vs. extremistas», *Dilemata: International Journal of Applied Ethics*, 34, 51–69.

*** Professor and Head of the Department of Philosophy I, University of Granada (Spain). Email: nef@ugr.es. Most of my work focuses on the applications of the philosophy of language to classical questions in the history of philosophy and to political and social practical concerns. Recent publications: Almagro, N., Hannikainen, I. & Villanueva, N. (forthcoming), «Whose words hurt? Contextual determinants of offensive speech», *Personality and Social Psychology Bulletin*, and Almagro, M. & Osorio, J. (forthcoming), «Weaponized testimonial injustice», *Las Torres de Lucca*.

**** This is a wholly collaborative paper, so order of authorship is arbitrary.

1. Introduction

> Despite the fact that you have the same sabbatical schedule as everyone else, he says,
> you are always on sabbatical. You are friends so you respond, easy.
> What do you mean?
> Exactly, what do you mean? (Rankine, 2014, 52)

At first sight, the previous exchange does not seem to count as one where the speaker is being offensive: the sentence "You are always on sabbatical" seems to aim at a *neutral* possible situation, which can be or not be the case. But as Rankine asks: exactly, what is the meaning communicated by the speaker through their claim in the above linguistic exchange? Does it really count as a simple description of how things are? Is it just a false description? The claim appears to mean something else. Is the speaker making an evaluation? Is it a joke? Is what has been said offensive in any way? Exactly, what does her friend mean? How can we know it? More context is needed to decide. But in what sense is the meaning we convey through our words context-sensitive?

The purpose of this paper is to explore the boundaries of a subset of the evaluative uses of language: offensive speech. More in particular, our goal is twofold. On the one hand, we want to chart the relationship between context and evaluative utterances, as it stands in the current literature. Forcefully, we will highlight certain recent findings, and say way less than it is necessary about other important phenomena. We will also focus, on the other hand, on the experimental study of a particular interaction between contextual information and our evaluative claims—when the context is able to turn a seemingly descriptive utterance into an evaluative one. For this second purpose, we will argue, certain recent positive proposals, in spite of their merit, come a bit short.

Evaluative uses of language seem to be importantly different from descriptive ones concerning the information they convey, despite the fact that the sentences commonly uttered to make claims from both language domains are structurally similar, even identical sometimes. We will highlight four different traits to showcase this difference: what is expressed, commonality of standards, retraction and practicality. First, descriptive claims seem to be embedded in practices where there is a high level of agreement in relation to how our surroundings are, while evaluative ones often exhibit lesser prior agreement on it. If I say that my room's door is closed, I'm simply describing how things around me are in an objective way, so to speak. But if I say that my room's door is *cool*, I'm also *expressing myself*, and my interlocutor might have a very different appreciation about the door (Cepollaro, Soria & Stojanovic, 2021) or, more importantly, about what a *cool door* amounts to. Second, descriptive and evaluative claims seem to give rise to very different situations of disagreement. If you and I disagree on whether the door of my room is closed, when the claim is just descriptive, then we seem to agree from the outset on how the dispute can be settled. In most scenarios, speaker and audience share common assumptions on how the claim could be verified. If they do not in fact share these assumptions, the situation stops making sense. If I suddenly discover that your preferred and only way to settle whether the door is closed is by asking the crystal skull that hangs from your necklace, we're

not involved in a discussion about our surroundings anymore (perhaps in one about the shortcomings of your upbringing, or none at all). This happens far less in situations where the disagreement revolves around an evaluative claim (Field, 2009, 2018). It's possible to discover that your musical standards are radically different from mine, and still continue discussing the relative greatness of Gesualdo and Palestrina. Third, retraction—understood as a linguistic move aimed at reassessing, from the actual context, a claim that we took to be true in the past—seems to work differently in descriptive and evaluative contexts: we demand that someone retracts evaluative statements more often than we do descriptive ones, while, perhaps surprisingly, it seems to be much more difficult to retract the former than the latter (Bordonaba & Villanueva, forthcoming). Finally, and most importantly for us here, evaluative claims exhibit a closer connection to action. If I say that my room's door is cool, I'm presenting myself as someone from whom certain courses of action can be reasonably expected, while claiming that the door is closed is not particularly action-guiding when it just conveys information about our surroundings (Gibbard, 2012, 44).

Of these four features, perhaps retraction and practicality require a bit more attention, given the pivotal role that they have played in the literature. Evaluative utterances are often assumed to carry an action-guiding component under their sleeve (see, for instance, Williams, 1985, 140; Gibbard, 1990, 33; Finlay, 2004; Strandberg, 2012, 89; Charlow, 2014; Chrisman, 2017, 27)—they provide reasons for action, they are conceptually linked to ulterior courses of action. We do not check and discover that our evaluative utterances were false, as it might be the case with descriptive utterances. I'm not simply wrong when I fail to live up to the courses of action that can be reasonably expected from my evaluations, I might be deemed irrational, a hypocrite, or someone who lacks the necessary linguistic competence. This, in turn, affects my ability to effectively retract my previous evaluations. By expressing affective attitudes, by making evaluations, I'm often inviting my audience to adopt a view on how things should or should not be that can have practical implications for those living around us. Harm can be inflicted, prejudices can be reinforced, in a way that it's difficult to counter whenever I change my mind and want to take back what I said. This is not to say that nothing can be done about our previous evaluative utterances whenever we change our mind and feel the urge to say so, we just need to be aware that making things right might go beyond simply saying that we no longer think what we previously thought.

As we will see, much attention has been paid to the context-dependent dynamics of evaluative language uses, more specifically to the contextual factors that might determine that a given object is positively or negatively evaluated. That is, how the parameters that determine our appreciation of an object might change from person to person, from time to time, from situation to situation. Much less attention has been devoted, however, to the contextual cues that determine that a piece of speech is interpreted evaluatively, rather than as mere description of our surroundings—whether or not a piece of speech is able to express the kind of affective attitudes that might turn out to be offensive, to contribute to the discrimination of disenfranchised groups. Only evaluative uses of language can turn out to be offensive, and only offensive pieces of discourse can express hatred, even if not every evaluation is offensive, and not every offense expresses hatred. We will explore in this paper both the context-dependency of those parameters that might make an evaluation offensive, and those

that contribute to the evaluative—and potentially offensive—evaluation of a seemingly neutral utterance, where a descriptive interpretation is also available.

Amongst the recent studies that have tackled the question of whether or not a given piece of discourse is offensive, and why, the role of context has been brought to the fore, at least at two different levels. Firstly, status and speakers' motives have been pointed out as deciding factors as to whether a single piece of discourse is offensive or it's simply a neutral description of a given situation (Gibson et al., 2019). Secondly, the syntactic position of the words has been highlighted as one of the factors that might determine whether a word is interpreted as a slur. This line has given rise to unexpected consequences (Cepollaro, Sulpizio & Bianchi, 2019): slurs are perceived as more offensive than other kinds of insults when they are presented in isolation. However, expressions of the form "Someone is an X" turn out to be perceived to be *less offensive* when X is a slur than when it is a non-slurring insult. Not only that, the offensive character of certain expressions only diminishes in indirect speech, in expressions of the form "Bilal believes that Lois is X", but doesn't disappear. Why should *I* be deemed as offensive for reporting on the thoughts of somebody else? To complicate matters even further, it remains an open question whether quoted slurs are potentially harmful, and thus rightfully deemed offensive, or not (Cepollaro & Zeman, 2020; Moreno & Pérez-Navarro, manuscript).

The studies just mentioned show how to go beyond the truism that offensiveness is context-dependent. Claiming that whether our utterances are offensive or not depends on the context is by now less of a claim than it is a cue to provide more relevant information.[1] This being so, incidentally, can provide a reasonable basis to derive a conversational implicature out of those cases in which supposedly knowledgeable people make that very same claim, as if it was all the explanation that a certain action needs to be understood. Unlike the examples in the studies above, our goal in this paper is not to examine the behavior of terms or phrases that are taken to be offensive by default, but rather to ask about expressions which are not usually identified as offensive, those that might ignite the most stubborn conversations about the nature of language, the intentions of the speakers, and so forth. We will review some of the contextual features that might determine how and when an otherwise perfectly descriptive expression can be used to express affective attitudes, to contribute to the stereotypes that harm disenfranchised minorities. One of the relevant determinants of offensive speech, as we will see, seems to be the speaker's status. If we turn back to Rankine's example knowing that the speaker is a white person and the target person is not white, our initial intuitions regarding the meaning communicated will probably vary. We will also, in § 3, explain why we should care about these examples at all. Before any of that, more has to be said about the context-dependent nature of evaluative language (§ 2).

2. Offensiveness and context-dependence

A friend argues that Americans battle between the "historical self" and the "self self." By this she means you mostly interact as friends with mutual interest and, for the most part, compatible personalities; however, sometimes your historical selves,

---

1   Cryptic? True. Potentially offensive? Also true.

her white self and your black self, or your white self and her black self, arrive with
the full force of your American positioning. Then you are standing face-to-face in
seconds that wipe the affable smiles right from your mouths. What did you say? Ins-
tantaneously your attachment seems fragile, tenuous, subject to any transgression of
your historical self. And though your joined personal histories are supposed to save
you from misunderstandings, they usually cause you to understand all too well what
is meant. (Rankine, 2014, 20)

The purpose of this section is to review some contributions that emphasize, in different
ways, the context-dependent nature of the evaluative. We will start by making some remarks
about the context-dependence of evaluative judgments. Next, we will review some linguistic
tests that, even though they are not exhaustive, point in the direction of the radical diffe-
rence between the descriptive and the evaluative in terms of their context-dependence. This
second type of context-dependency is slightly different: evaluative judgments are highly
context-sensitive, but determining whether a judgment is descriptive or evaluative, a prior
step, is also highly context-dependent. Finally, we will focus on the influence exerted by
different contextual factors in the process of determining whether the meaning of a claim is
descriptive or evaluative/offensive.

First, evaluative uses of language are context-dependent in at least two different senses.
One of the characterizing features of evaluative judgments seems to be its gradable and
multidimensional nature. As McNally and Stojanovic highlighted, when we evaluate an
object, an event, a situation, a person, etc., the evaluation can come in degrees, and we
might hinge on different features of the context constituting the "threshold of applicability"
(McNally & Stojanovic, 2017, 21). Evaluative adjectives—adjectives whose use by default
is evaluative—are gradable and multidimensional. For instance, when evaluating whether a
natural enclave is beautiful or whether someone is healthy, we can say that it is *more* or *less*
beautiful or healthy. Moreover, we can give consideration to the regularity of the terrain, the
sort of vegetation, the color of the sky, the number of things found there, etc., or to the state
of their cardiovascular system, immune system, habits, cholesterol, etc. Thus, the gradability
of an expression requires that a parameter is contextually fixed, and the parameter might
vary from one speaker to another, from one situation to the next. Its multidimensionality,
for its part, is characterized by a contextually-sensitive list of features that determine the
threshold, the bar that an object has to jump over, so to speak, in order to receive positive
or negative valence. In this sense, the features that determine whether an evaluation is going
to be positive or negative depend on the context.

Furthermore, the relative weight of each of the dimensions involved in setting the thres-
hold might vary from context to context (McNally & Stojanovic, 2017; Cepollaro, Soria &
Stojanovic, 2021). For instance, I might take into consideration the regularity of the terrain
much more than any other parameter when deciding which place is more beautiful to take
a walk, but give more weight to the color of the sky when deciding which place is more
beautiful to lay down and watch the clouds passing by. And I might give much more weight
to a person's cardiovascular state when deciding who is healthy enough to play on a sports
team, but give less value to that parameter in another situation. It might even be the case that

the relative weight of each dimension varies when considering two different places while deciding which one is more beautiful for a walk.

Second, a number of theories have accommodated value-related information within the realm of pragmatics, a highly context-dependent layer of meaning (see, for instance, Corredor, 2014). This realm includes approaches that place certain contextual information beyond what we literally say, in the range of what we implicate or presuppose by saying what we say. That is, our utterances can convey information that isn't part of the literal meaning of the sentences we utter, and do so by exploiting certain linguistic mechanisms and conversational principles. The main theoretical notions within this strand are *presuppositions* and different types of *implicatures*: conventional, generalized conversational and particularized conversational implicatures.

Several authors have defended a presuppositional account of different sorts of evaluative judgments and expressions (see, for example, Macià, 2002; Cepollaro, 2015, 2016, 2017a, 2017b; Cepollaro & Stojanovic, 2016). The main point of this approach is that it situates the evaluative meaning at the level of what is taken for granted. For instance, if I evaluate Loyiso Gola's special *Unlearning* by saying that it is *powerful*, I might literally mean that the show has a strong effect on people's thought, and trigger the evaluative presupposition that things that have a strong effect on people's thought are good precisely because they do so. That is, our evaluations, at least in some cases, implicitly take for granted a certain perspective, a certain set of beliefs concerning what is good and what is bad, which is situated at the level of what is presupposed.

Other authors situate the meaning conveyed by evaluative uses of language at the level of what is conventionally implicated (see, for example, Stenner, 1981; Potts, 2005, 2007; Copp, 2009; Whiting, 2007, 2013; Williamson, 2009; McCready, 2010; Gutzman, 2011). Implicatures are pieces of information conveyed by what we say, that is, information that goes beyond the literal meaning expressed through an utterance. Some implicatures are part of the conventional meaning of what is literally said. As Abbot claims, presuppositions and conventional implicatures are "close neighbors" (Abbott, 2006, 2). In contrast to presuppositions, conventionally implicated information is not taken for granted but inferentially conveyed. For instance, if I evaluate Loyiso Gola's show *Unlearning* by saying that it is powerful, I conventionally implicated the proposition that Loyiso Gola's show is good, given the conventional meaning of "powerful".

The two further types of implicatures, particularized and generalized conversational ones, are implicated pieces of information that exploit conversational rules. They allow us to restore the conversational order when some basic rule is apparently bent. The main difference between them is their degree of context-dependence. Particularized implicatures require some special contextual features, while generalized ones arise in the absence of particular circumstances that would defeat it (see Väyrynen, 2013). As an example of a conversational implicature, take the utterance of the sentence "It lasts more than an hour" as a response to the question "How was the show?" to implicate that it was not good.

As we have seen, the specific valence of our evaluations depends on the context in at least two ways: the features that determine the positive or negative valence of a judgment depend on the context, and the relative weight of those features might also vary from context to context. Moreover, a good number of theories have accommodated the

information conveyed through our evaluative judgments within the realm of pragmatics, a highly context-dependent layer of meaning. When a judgment is identified as an evaluative one, then it is highly context-dependent, in different senses. However, there is a third issue that might also need the intervention of context. By way of example: dogwhistles. Dogwhistles are a sort of political manipulation that some carry out to covertly obtain certain political benefits. Succinctly, they are pieces of speech that convey certain information which is consciously or unconsciously received just by a subset of the general audience (Saul, 2018; Torices, 2021). Only under certain very specific circumstances can a piece of public discourse be interpreted as a dogwhistle, taken to express affective attitudes only to a given proper subset of the target audience, while remaining "neutral" to the rest of it. Only with respect to this subset of the audience is it able to contribute to the promotion of negative stereotypes, to be offensive. This selective expression of hatred carried out by dogwhistling is also highly context-dependent. Note, however, that context plays a role here not only once the judgment is identified as an evaluative one, but in a previous step. The claim will express affective attitudes or it will not, will be an evaluative claim or it will not, depending on the context.

So, there are linguistic phenomena characterized by promoting certain affective attitudes in a subset of the general audience whose success is also highly context-dependent. In this sense, the role played by the context in the determination of the meaning conveyed through an utterance generalizes a bit further. Determining whether a piece of discourse is evaluative or descriptive, potentially offensive or "neutral", also requires taking into account several features of the context. At the beginning of this paper, we have already introduced some intuitive and powerful tests to differentiate a factual claim from an evaluative one: what is expressed, commonality of standards, retraction and practicality. Let's add two more tests to be able to spot gradable and multidimensional predicates, the kind of expressions that are typically involved in evaluative claims. A descriptive claim such as "Gola's show was recorded at 3 PM" does not admit of degrees—e.g., "# Gola's show was *very / slightly / more* recorded at 3 PM than the previous one". Neither it is possible to distinguish different respects playing a role—e.g., "# Gola's show was recorded at 3PM *in every / some respect / with respect to*". However, an evaluative claim such as "Gola's show was nice" does admit degrees—e.g., "Gola's show was *very / slightly* nicer than the previous one—and multidimensionality—e.g., Gola's show was nice *in every / some respect / with respect to* the performance, but not with respect to the plot (Cepollaro, Soria & Stojanovic 2021). Being gradable is a cue of being evaluative, though it does not warrant it. Not every gradable and multidimensional adjective, or predicative expression, is used to make evaluations, and not every evaluation is made with the aid of these expressions. Firstly, any multidimensional and gradable adjective can be used descriptively in a context in which the standards are fixed. Secondly, given the appropriate context, almost any predicative expression can be used to make an evaluation, once it is established that possessing that particular feature is associated with a valence, positive or negative.

Soria and Stojanovic have distinguished some additional tests (Soria & Stojanovic, 2019). The first one is the *juxtaposition with "although"-type connectives* test, based on the action-guidance of the evaluative. The idea is that a complex sentence composed of two sentences connected by the conjunction "although", establishing a contrast between

a claim and a practical attitude, seems felicitous only when the claim is evaluative. For instance, the evaluative claim "Nassim thinks that eating meat is wrong, although she consumes twice more meat than average" seems felicitous, while the descriptive claim "Nassim thinks that eating meat is unpopular, although she consumes twice more meat than average" does not. Another test is the *lack of epistemic justification*. The idea is that specific evaluative judgments, such as claiming "I've just watched Gola's show and it's amazing", won't admit the question "How do you know?", while descriptive claims such as "Gola's show lasts one hour and a half" does. Finally, we also have the *lack of lying potential* test. Evaluative claims do not admit the reply "That's a lie", in part because they are a genuine expression of our own perspective. For instance, the claim "This show is amazing" cannot be replied by saying "That's a lie", while a descriptive one such as "This show is the second one of her career" can. As shown above with the first two features highlighted in this linguistic road to the evaluative, these tests also fall short from being conclusive. It's perfectly OK to inquire into the reasons for a given evaluation, especially if the topic is important to us, or we are close to the speaker. "The White Album better than *Abbey Road*? I can't believe you said that!", "How can you take A over B?", "The White Album better than *Abbey Road*? Are you sure about that?" all seem perfectly fine to us, evaluative claims where the grounds are explicitly questioned. Seemingly, instances can be found of evaluative claims followed by "That's a lie". Imagine you keep telling a child for years "Eat your beans, they're delicious". Years later, the now-adolescent discovers that you'd rather skip a meal than eating beans and protests: "You hated beans, and yet you kept telling me they were delicious, you were lying to me the whole time!!".

The first of the tests mentioned above requires a bit more consideration. It hinges on the particular connection with our practical commitments that evaluative utterances seem to have, where the uptake seems to change the practical commitments of participants in the conversation. This rings true enough, it looks reasonable to assume that if I'm taking part in a conversation, and I let an evaluative claim be made unchallenged, the speaker has every reason to suppose that I *also* believe what they expressed. If we do not react to racist claims, we are contributing to the normalization of these words, and potentially reinforcing negative stereotypes. On the other hand, it also looks a bit excessive as a requirement for evaluative uses of language. A perfectly evaluative utterance can be understood, and be given a *pass*, even if we disagree with the claim, or with the way it is expressed. It's too strong to demand from every participant in a conversation the attitude of a linguistic *vigilante*, ready at all times to deny the uptake to evaluative utterances, whenever there's a disagreement. Sometimes staying quiet is required to let others find their way on their own. Speakers need to be given the chance to recognize the attitudes that they express, even if they were not aware of them before making their utterances, and this is much easier, in most contexts, if we do not adopt the attitude of linguistic *vigilantes*. Again, Claudia Rankine captures both intuitions beautifully in the following passage:

> Another friend tells you you have to learn not to absorb the world. She says some-
> times she can hear her own voice saying silently to whomever—you are saying
> this thing and I am not going to accept it. Your friend refuses to carry what doesn't
> belong to her.

> You take in things you don't want all the time. The second you hear or see some ordinary moment, all its intended targets, all the meanings behind the retreating seconds, as far as you are able to see, come into focus. Hold up, did you just hear, did you just say, did you just see, did you just do that? Then the voice in your head silently tells you to take your foot off your throat because just getting along shouldn't be an ambition. (Rankine, 2014, 55)

All these tests, even if they are not always completely effective, showcase the difference between the descriptive and the evaluative. But, beyond the features that the evaluative exhibits so far, such as the difference in what is expressed, commonality of standards, retraction and practicality; beyond the context-dependence, in different senses, of evaluative judgments, and beyond the linguistic particularities that are sometimes exhibited by the expressions that we use to evaluate, there is more that we need to know concerning *when* a seemingly neutral statement is indeed used to express affective attitudes. This is also highly context-dependent. The relationship between contextual information and our linguistic behavior conducive to conveying evaluations is a complex one, and the limitations of the criteria pointed above tell this story. Our purpose in this paper, as stated in the introduction, is to chart a conceptual map of the territory, exploring the benefits and shortcomings of the positive proposals contained in the recent literature, but *also* to point to a different territory, one that involves experimental philosophy and that will require more attention in the near future.

Recently, we have conducted two empirical studies on the influence of different contextual factors in assessing the same statement as simply informative or also offensive (Almagro, Hannikainen & Villanueva, 2021). In this research, we took into consideration different contextual features. In particular, we designed a first study with a set of 8 vignettes where a speaker utters a seemingly descriptive sentence, and manipulated three contextual factors: speaker status, speaker intention, and harm caused to the audience. All participants viewed eight vignettes telling a short story in which a speaker claims something, and each participant was exposed to a randomized combination of factors in each vignette. Then, we asked each participant to rate from 1 (strongly disagree) to 7 (strongly agree) the following claims after reading each scenario: "The speaker simply offers information"; "The speaker is being offensive". We added an additional question about the acceptability of the claim with two available options: "The speaker shouldn't say that kind of thing" / "I don't see any problem in the speaker saying that kind of thing". Our results showed that for participants, although the three factors' effect was significant, the speaker's status played a more important role than the intention and the outcome factors in taking the same statement as offensive and prohibited (in a second study we have disambiguated between speakers' membership and social status, and membership appears to exert a greater effect). So, according to these findings, all three factors are relevant, and speaker's status and membership are even more important contextual factors in determining whether a claim expresses affective attitudes or just conveys descriptive information. Moreover, status seemed to play a more important role than agents were willing to explicitly acknowledge, when asked directly. Thus, the context does not only affect our evaluations because our standards are diverse and might change from situation to situation; whether some utterance is deemed as evaluative—and potentially offensive—or not can only be determined by paying attention to features of the

context such as: who's speaking, what are the potential consequences of the utterance, and what intention is reasonable to attribute to the speaker.

Let's turn now to the initial example of this section. What did the speaker say by claiming that Americans battle between the "historical self" and the "self self"? Was it good or bad? Is what they said offensive? The moral of this section is that before examining the context to check what standards are in play, we need to consider the actual expressions used and their potential occurrences in other linguistic contexts, but also the status of the speaker and the audience, their intentions, and the possible consequences.

3. Why should we care?

> A woman you do not know wants to join you for lunch. You are visiting her campus. In the café you both order the Caesar salad. This overlap is not the beginning of anything because she immediately points out that she, her father, her grandfather, and you, all attended the same college. She wanted her son to go there as well, but because of affirmative action or minority something—she is not sure what they are calling it these days and weren't they supposed to get rid of it?—her son wasn't accepted. You are not sure if you are meant to apologize for this failure of your alma mater's legacy program; instead you ask where he ended up. The prestigious school she mentions doesn't seem to assuage her irritation. This exchange, in effect, ends your lunch. The salads arrive. (Rankine, 2014, 19)

Although it might appear to be a simple description, saying to a non-white person that you wanted your son to go to her college but that he wasn't accepted because of affirmative action or minority something *is* offensive, even if you didn't intend it that way. Take the following case. In 2008, Phil Edwards and Ian Roberts published a short piece in the British peer-reviewed general medical journal *The Lancet*. In this article, the authors tackled the issue of global food insecurity and climate change. In particular, they made the point that reducing the demand for transportation fuel is important for reducing food prices in two senses (Edwards & Roberts, 2008). The first argument offered by the authors is as follows: motorized transport is more than 95% oil-dependent. Since oil is a key agricultural input, demand for transportation fuel increases its price, and then affects food prices. The second argument goes like this: increased car use contributes to rising food prices by promoting obesity, which increases the global demand for food. Additionally, more transportation fuel energy will be used to transport the increased mass of the obese population. Their arguments were presented under a descriptive guise:

> We estimate that a population of 1 billion people with a stable mean body-mass index (BMI) of 24·5 kg/m2 consumes an average 6·5 MJ of food energy per person per day to maintain basal metabolic rate, and a further 4 MJ per person per day for activities of daily living. An obese population of 1 billion people with a stable mean BMI of 29·0 kg/m2 would require an average 7 MJ of food energy per person per day to maintain basal metabolic rate, and 5·4 MJ per person per day for activities of daily

> living. Compared with the normal weight population, the obese population consumes 18% more food energy. (Edwards & Roberts, 2008)

From this, the authors concluded that "urban transport policies that promote walking and cycling would reduce food prices by reducing the global demand for oil, and promotion of a normal distribution of BMI would reduce the global demand for, and thus the price of, food" (Edwards & Roberts, 2008). Although this contribution might be initially seen just as a descriptive one, and it is presumably written with that intention, some media reports echoed Edwards and Roberts' findings under headlines such as "Obese *blamed* for the world's ills". In the same journal, Harald Schmidt wrote a reply to the authors claiming that Edwards and Roberts' discussion of the role of overweight people in climate change is *ethically problematic* (Schmidt, 2008). He claims that some important alternatives to the authors' hypothesis, such as cutting unnecessary transport, using alternative fuels, lighter materials in vehicles, etc., are neither considered nor acknowledged, and that it makes it more problematic to advocate that reducing weight is a key priority in fighting against climate change and food prices. So, their contribution, he concludes, counts as an instance of victim blaming of an already marginalized and vulnerable population.

In addition to the authors' narrow analysis that leads to focus on a single possible explanation that, as a result, blames certain people from a disenfranchised group, there are other rather problematic assumptions in play. For instance, Edwards and Roberts assume that someone's body-mass is always proportional to the intake of a particular amount of food energy, which is a stereotype. It is also assumed that there is a *normal* body-mass index, which makes other body-mass indexes of healthy people seem abnormal. Empirical studies have shown that the use of terms such as "obese", "fat", etc., as "descriptive" as they might look in a scientific context, promotes negative attitudes towards people to whom such predicates are applied (see, for instance, Brochu & Esses, 2011).

Nevertheless, it can be convincingly argued that doing all this was not the authors' intention. Presumably, the authors' purpose was simply to genuinely make a contribution to the topic, and not to blame anyone. However, the meaning of our words is beyond our acknowledged intentions, and we seem to be especially bad at knowing our own minds. For example, we might be completely honest when we say that we consider all races to be equal, and yet still display racist attitudes in our linguistic and nonlinguistic behavior. The evaluative nature of the information we convey through our verbal and nonverbal actions gives us clues about our own personality, what to keep and what to try and keep in check.

In our study presented above, we also found some surprising results along this line. At the end of the experiment, we included a general final question in which we explicitly asked about the influence of the three factors (speaker's status, intention, and the outcome) when considering the same statement as offensive or not. Participants had to rate from 1 (little influence) to 7 (much influence) speaker status, speaker intention, and the harm felt by the audience. Although participants had shown through their concrete case-by-case responses that the identity of the speaker, both in terms of membership and social status, is the contextual factor that most affected their responses in considering the same statement as offensive or neutral, when approached in an abstract way, participants rated the speaker's intention as the most relevant factor, and the speaker's identity as the least important. So,

it can be concluded that most of the participants in our study sincerely say that they think that the speaker's intention is the most relevant contextual factor in considering a claim as descriptive or offensive, but that they don't actually think so, given their responses in specific situations. This result is in line with other research suggesting that we have systematic blindspots regarding our own mental states (Nisbett & Wilson, 1977; Zajonc, 2001; Wilson, 2002; Schwitzgebel, 2008, 2011; Carruthers, 2011). Our evaluations can make us realize an aspect of our personality, of our mind, that we could have otherwise ignored.

Perhaps this realization comes a bit too late, only when we catch ourselves promoting the kind of stereotypes that reinforce the discrimination against disenfranchised groups, we do acknowledge that perhaps there was an aspect of our personality that we needed to address. It's too late to take our utterance back, and the harm is probably already done, but it might be one of the very few ways to access those dispositions of our own, and we'd better keep them in mind. Take an implicit test, trust those that are close to you, pay attention to what you say, perhaps inadvertently, especially when you are called out on that. What did you say? Exactly, what did you mean? Besides giving information about our own mind, evaluations promote negative attitudes and have other pernicious consequences, especially when they are unnoticeably made or presented under a descriptive guise. For instance, Samuel Huntington's contribution to the Report on the Governability of Democracies to the Trilateral Commission, published in 1975, was to argue that the problems of the United States in the 1970s result from "an excess of democracy". His recommendation for the United States, as Stanley puts it, was "to try to reinstall some measure of obedience to authority by making various central domains in life, ones that should be governed democratically, the domain of experts" (Stanley, 2015, 40). That is, Huntington was recommending to take certain matters of value as if they were factual ones. Appealing to experts as if they have the last word when deciding about a normative issue is covertly making an evaluation, presenting it as if it were a description. In doing so, one may appear to be making a reasonable contribution when in fact one is undermining certain democratic ideals as Stanley points out. The subtle effect exerted by different contextual factors on the meaning conveyed through a claim, together with our inability to clearly identify the influence exerted by these factors, makes us especially vulnerable to situations such as the previous ones. If Edwards and Roberts say that they didn't intend to be offensive, then someone might conclude erroneously that what they said is not offensive, and that it counts as a reasonable contribution. Similarly, one might wrongly conclude that, since Huntington appealed to experts, he was talking about a factual issue, and that he was making a reasonable contribution.

This move can be done to advance certain political agendas. As Kappel puts it, "in some cases, there might be strategic advantages to framing a disagreement as fact-dependent, even when it is not. Or conversely, disagreements may be framed as value-dependent, when they are really fact dependent, and so on" (Kappel, 2017, 317). In particular, this move generates situations of *crossed disagreements* (Osorio & Villanueva, 2019). A crossed disagreement is a situation in which each of the parties of a genuine disagreement displays clear signs of conceiving the dispute in significantly different terms, e.g., one conceives it in descriptive terms, and the other as if it were a normative dispute. One of the pernicious consequences of the systematic confusion of the descriptive and the evaluative in public debates is that they foster polarization: crossed disagreements expand the pool of arguments supporting

the views held by each party (Osorio & Villanueva, 2019; Almagro, Osorio & Villanueva, 2021), which in turn contributed to the raise of polarization. In other words, increasing the size and density of the pool of arguments that one is exposed to fosters polarization: when we are exposed to a limited group of arguments just in favor of one position, especially when they are frequently repeated, we tend to become more impervious to others' reasons (Sunstein, 2017; Barberá et al., 2015; Vicario et al., 2016). Since the rise of polarization increases distrust in public institutions and governments, it makes citizens less likely to seek different perspectives on controversial topics, and corrodes the proper functioning of democratic institutions, among other things (see Levitsky & Ziblatt, 2018; Carothers & O'Donohue, 2019), the systematic confusion of the descriptive and the normative in public debates is pretty dangerous.

In this paper we've taken a much more personal approach, focusing on cases where the speakers are expressing hatred through apparently descriptive claims. Let's briefly present other cases in this vein. In a scientific article, Simko and Ginter concluded that the uncontrolled sugar consumption leading to high rates of diabetes explains the low life expectancy of the Roma population in Europe (Simko & Ginter, 2010), ignoring many other relevant data, such as that Roma population regularly appears in the Eurobarometer as the minority to which the rest of the people in Europe feel less affinity, have less access to resources, suffer more poverty, etc. Their food choices might not be a result of their culture or their way of life, but the sheer consequence of their low income. A recent report on disparities in the risk and outcomes of COVID-19 in the United Kingdom (Public Health England, 2020), while it recognizes that non-white males have been more likely to die from a COVID-19-related condition than White males, highlights that "an analysis of over 10,000 patients with COVID-19 admitted to intensive care in UK hospitals suggests that, once age, sex, obesity and comorbidities are taken into account, there is no difference in the likelihood of being admitted to intensive care or of dying between ethnic groups", and adds that people of non-white groups "are also more likely than people of White British ethnicity to be born abroad, which means they may face additional barriers in accessing services that are created by, for example, cultural and language differences". In doing so, again, they come short of blaming people from some disenfranchised groups for their bad fortune, which is highly offensive and promotes negative attitudes towards them, despite the authors neither having this intention nor using a language that is clearly evaluative.

4. Conclusion

In this paper, we have explored the boundaries of offensive speech, a subset of evaluative uses of language. In particular, we have focused on the contextual dependence of offensive speech: the contextual cues that determine that a piece of speech is interpreted evaluatively, rather than as mere description of our surroundings, and is thus able to express the kind of affective attitudes that might contribute to the discrimination of disenfranchised groups. First, we have dwelt on how context affects evaluative meaning. As we have seen, our evaluations depend on a variety of contextual features, and the relative weight of those features also varies from context to context. Second, we have briefly reviewed some approaches to the evaluative from the realm of pragmatics, the classic

locus of context-dependence. Then, we have focused on a more general issue: how the context plays a role in determining whether a piece of speech expresses affective attitudes or remains neutral. We have revisited different tests and remarks that can help us to know whether a claim is descriptive or evaluative, and we have introduced an empirical study that we have recently conducted on the influence exerted by three contextual factors in determining whether a statement counts as a descriptive or evaluative one: speaker's status, speaker's intention and outcome. Finally, we have discussed several cases where the speaker's claim seems to be descriptive but it actually expresses affective attitudes towards certain groups of people, and showed why they are problematic and why we need to be able to recognize when a claim is an evaluative one.

The goal of this paper started from the need of making explicit certain forms of injustice that frequently go unnoticed, in order to be able to fight against them. In this sense, it belongs to *the political turn in analytic philosophy* (Pinedo & Villanueva, forthcoming). The defining feature of this political turn is taking our raw perceptions of injustice as the starting point, and bringing to the fore the capacity of our philosophical tools in detecting and intervening on oppressive and unjust practices. Our best theories will be the ones that more efficiently serve this purpose: identifying and intervening in certain social injustices.

> Words work as release—well-oiled doors opening and closing between intention, gesture. A pulse in a neck, the shiftiness of the hands, an unconscious blink, the conversations you have with your eyes translate everything and nothing. What will be needed, what goes unfelt, unsaid—what has been duplicated, redacted here, redacted there, altered to hide or disguise—words encoding the bodies they cover. And despite everything the body remains.
> Occasionally it is interesting to think about the outburst if you would just cry out—To know what you'll sound like is worth noting. (Rankine, 2014, 69)

References

Abbott, B. (2006), "Where have some of the presuppositions gone", in B. Birner & G. Ward, (eds.), *Drawing the Boundaries of Meaning: Neo-Gricean Studies in Pragmatics and Semantics in Honor of Laurence R. Horn*, Amsterdam: Studies in Language Companion Series, pp. 1–20.

Almagro, M., Hannikainen, I. & Villanueva, N. (2021), "Whose words hurt? Contextual determinants of offensive speech", *Personality and Social Psychology Bulletin*.

Almagro, M., Osorio, J. & Villanueva, N. (2021). "Weaponized testimonial injustice", *Las Torres de Lucca*.

Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A. & Bonneau, R. (2015), "Tweeting from left to right: Is online political communication more than an echo chamber?", *Psychological Science*, 26(10), 1531–1542.

Bordonaba, D. & Villanueva, N. (forthcoming). "Retractación y contextualismo: nuevas condiciones de adecuación", in D. Pérez Chico (ed.), *Contextualismo semántico*, Zaragoza: Prensas de la Universidad de Zaragoza.

Brochu, P. M., & Esses, V. M. (2011). "What's in a name? The effects of the labels "fat" versus "overweight" on weight bias", *Journal of Applied Social Psychology*, 41(8), 1981–2008.

Carothers, T., & O'Donohue, A. (2019), *Democracies Divided: The Global Challenge of Political Polarization*, Washington, DC: Brookings Institution Press.

Carruthers, P. (2011), *The Opacity of Mind: An Integrative Theory of Self-Knowledge*, Oxford: Oxford University Press.

Cepollaro, B. (2015), "In defense of a presuppositional account of slurs", *Language Sciences*, 52, 36–45.

Cepollaro, B. (2016), "Building evaluation into language", *Phenomenology and Mind*, 11, 158–168.

Cepollaro, B. (2017a), "The shortcut of discrimination", *Rivista di Estetica*, 64, 53–65.

Cepollaro, B. (2017b), *The Semantics and Pragmatics of Slurs and Thick Terms*, PhD thesis, Paris: PSL Research University.

Cepollaro, B. & Stojanovic, I. (2016), "Hybrid evaluatives: In defense of a presuppositional account", *Grazer Philosophische Studien*, 93(3), 458–488.

Cepollaro, B., Sulpizio, S. & Bianchi, C. (2019), "How bad is it to report a slur? An empirical investigation", *Journal of Pragmatics*, 146, 32–42.

Cepollaro, B. & Zeman, D. (2020), "The challenge from non-derogatory uses of slurs", *Grazer Philosophische Studien*, 97(1), 1–10.

Cepollaro, B., Soria, A. & Stojanovic, I. (2021), "The semantics and pragmatics of value judgments", in P. Stalmaszczyk (ed.), *The Cambridge Handbook of the Philosophy of Language*, Cambridge: Cambridge University Press.

Charlow, N. (2014), "The problem with the Frege-Geach problem", *Philosophical Studies*, 167(3), 635–665.

Chrisman, M. (2016), *What Is This Thing Called Metaethics*, London: Routledge.

Corredor, C. (2014), "Pejoratives and social interaction", in P. Stalmaszczyk (ed.), *Issues in Philosophy of Language and Linguistics*, Łódź: Wydawnictwa Uniwersytetu Łódzkiego, 39–55.

Copp, D. (2009), "Realist-expressivism and conventional implicature", *Oxford Studies in Metaethics*, 4, 167–202.

Edwards, P. & Roberts, I. (2008), "Finding long-term solutions to the world food crisis", *The Lancet*, 371(9622), 1389.

Field, H. (2009), "Epistemology without metaphysics", *Philosophical Studies*, 143(2), 249–290.

Field, H. (2018), "Epistemology from an evaluativist perspective", *Philosophers' Imprint*, 18(2), 1–23.

Finlay, S. (2004), "The conversational practicality of value judgment", *Journal of Ethics*, 8(3), 205–223.

Gibbard, A. (1990), *Wise Choices, Apt Feelings*, Cambridge: Cambridge University Press.

Gibbard, A. (2012), *Meaning and Normativity*, Oxford: Oxford University Press.

Gibson, J. L., Epstein, L. & Magarian, G. (2019), "Taming uncivil discourse", *Political Psychology*, 41(2), 383–401.

Gutzman, D. (2011), "Expressive modifiers & mixed expressives", *Empirical Issues in Syntax and Semantics*, 8, 123–141.

Kappel, K. (2017), "Fact-dependent policy disagreements and political legitimacy", *Ethical Theory and Moral Practice,* 20(2), 313–331

Levitsky, S., & Ziblatt, D. (2018), *How Democracies Die*, New York: Crown.

Macià, J. (2002), "Presuposición y significado expresivo", *Theoria: Revista de Teoría, Historia y Fundamentos de la Ciencia*, 3(45), 499–513.

McCready, E. (2010), "Varieties of conventional implicature", *Semantics and Pragmatics*, 3(8), 1–57.

McNally, L. & Stojanovic, I. (2017), "Aesthetic adjectives", in J. Young (ed.), *The Semantics of Aesthetic Judgment*, Oxford: Oxford University Press.

Moreno, A. & Pérez-Navarro, E. (manuscript), "Beyond the conversation: The pervasive danger of slurs".

Nisbett, R., & Wilson, T. (1977), "Telling more than we can know: Verbal reports on mental processes", *Psychological Review*, 84(3), 231–259.

Osorio, J. & Villanueva, N. (2019), "Expressivism and crossed disagreements", in M. J. Frápolli (ed.), *Expressivisms, Knowledge and Truth*, Cambridge: Cambridge University Press, pp. 111–132.

Pinedo, M. & Villanueva, N. (forthcoming), "Epistemic de-platforming", in D. Bordonaba, V. Fernández & J. R. Torices (eds.), *The Political Turn in Analytic Philosophy: Reflections on Social Injustice and Oppressions*, Berlin: De Gruyter.

Potts, C. (2005), *The Logic of Conventional Implicatures*, Oxford: Oxford University Press.

Public Health England (2020), "Disparities in the risk and outcomes of COVID-19", https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/908434/Disparities_in_the_risk_and_outcomes_of_COVID_August_2020_update.pdf

Rankine, C. (2014), *Citizen: An American Lyric*. Minneapolis, MN: Graywolf Press.

Saul, J. (2018), "Dogwhistles, political manipulation and philosophy of language", in D. Fogal, D. W. Harris & M. Moss (eds.), *New Works on Speech Acts*, Oxford: Oxford University Press, pp. 360–383.

Schmidt, H. (2008), "Transport policy, food policy, obese people, and victim blaming", *The Lancet*, 372(9622), 627.

Simko, V. & Ginter E. (2010), "Short life expectancy and metabolic syndrome in Romanies (gypsies) in Slovakia", *National Center for Biotechnology Information*, 18(1), 16–18.

Stanley, J. (2015), *How Propaganda Works*, Princeton, NJ: Princeton University Press.

Stenner, A. J. (1981), "A note on logical truth and non-sexist semantics", in M. Vetterling-Braggin, (ed.), *Sexist Language: A Modern Philosophical Analysis*, New York: Littlefield, Adams & Co., pp. 299–306.

Strandberg, C. (2012), "A dual aspect account of moral language", *Philosophy and Phenomenological Research*, 84(1), 87–122.

Sunstein, C. (2017), *#Republic: Divided Democracy in the Age of Social Media*, Princeton, NJ: Princeton University Press.

Schwitzgebel, E. (2008), "The unreliability of naïve introspection", *The Philosophical Review*, 117(2), 245–273.

Schwitzgebel, E. (2011), *Perplexities of Consciousness*, Cambridge, MA: MIT Press.

Torices, J. R. (2021). "Understanding dogwhistles politics", *Theoria. An International Journal for Theory, History and Foundations of Science*. https://doi.org/10.1387/theoria.22510

Väyrynen, P. (2013), *The Lewd, the Rude and the Nasty: A Study of Thick Concepts in Ethics*, Oxford: Oxford University Press.

Vicario, M. D., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., Stanley, H. E. & Quattrociocchi, W. (2016), "The spreading of misinformation online", *Proceedings of the National Academy of Sciences*, 113(3), pp. 554–559.

Whiting, D. (2007), "Inferentialism, representationalism and derogatory words", *International Journal of Philosophical Studies*, 15(2), 191–205.

Whiting, D. (2013), "It's not what you said, it's the way you said it: Slurs and conventional implicatures", *Analytic Philosophy*, 54(3), 364–377.

Williams, B. (1985), *Ethics and the Limits of Philosophy*, London: Fontana.

Williamson, T. (2009), "Reference, inference, and the semantics of pejoratives", in J. Almog & P. Leonardi (eds.), *The Philosophy of David Kaplan*, Oxford: Oxford University Press, pp. 137–159.

Wilson, T. (2002), *Strangers to Ourselves: Discovering the Adaptive Unconscious*, Cambridge, MA: Harvard University Press.

Zajonc, R. (2001), "Mere exposure: A gateway to the subliminal", *Current Directions in Psychological Science*, 10(6), 224–228.