

Beyond the Conversation: The Pervasive Danger of Slurs


Alba Moreno* – Eduardo Pérez-Navarro**

Received: 3 December 2020 / 1st Revised: 14 April 2021 /
2nd Revised: 31 May 2021 / Accepted: 9 August 2021

Abstract: Although slurs are conventionally defined as derogatory words, it has been widely noted that not all of their occurrences are derogatory. This may lead us to think that there are “innocent” occurrences of slurs, i.e., occurrences of slurs that are not harmful in any sense. The aim of this paper is to challenge this assumption. Our thesis is that slurs are always potentially harmful, even if some of their occurrences are nonderogatory. Our argument is the following. Derogatory occurrences of slurs are not characterized by their sharing any specific linguistic form; instead, they are those that take place in what we call uncontrolled contexts, that is, contexts in which we do not have enough knowledge of our audience to predict what the uptake of the utterance will be. Slurs uttered in controlled contexts, by contrast, may lack derogatory character. However, although the kind of context at which the utterance of a slur takes place can make it nonderogatory, it cannot completely deprive it of its harmful potential. Utterances of

* Universidad de Granada


 <https://orcid.org/0000-0002-6883-4357>

 Facultad de Psicología, Despacho 256, Campus Universitario de Cartuja, s/n
18011, Granada, Spain.

 almorenozurita@gmail.com

** Universidad de Granada

 <https://orcid.org/0000-0002-2240-2380>

 Facultad de Psicología, Despacho 256, Campus Universitario de Cartuja, s/n
18011, Granada, Spain.

 edperez@ugr.es



slurs in controlled contexts still contribute to normalizing their utterances in uncontrolled contexts, which makes nonderogatory occurrences of slurs potentially harmful too.

Keywords: Context; derogation; nonderogatory occurrences of slurs; normalization; slurs.

1. Introduction

Although slurs are conventionally defined as derogatory words, it has been widely noted that not all of their occurrences are derogatory. Cases of *mention*, in which we talk about the word rather than applying it to anybody, are the ones that most straightforwardly come to mind (see e.g. Hornsby 2001, 129–30). However, some full-fledged *uses* of slurs are standardly taken to be nonderogatory too. Among these, two kinds of uses have been most discussed. On the one hand, members of the target group can *appropriate* a slur in order to demarcate the group or foster solidarity or feelings of belonging, thus being able to use it in a nonderogatory way (see Bianchi 2014; Cepollaro 2017). But, on the other hand, we can also find nonderogatory uses of slurs that are not instances of appropriation—what have been called *nonderogatory, nonappropriated* (NDNA) uses of slurs (Hom 2008; see also Croom 2011, and section 2 of this paper for examples).¹ The fact that not all occurrences of slurs are derogatory, as mentions, appropriated and NDNA uses seem to prove, may lead us to think that there are “innocent” occurrences of slurs, i.e., occurrences of slurs that are not harmful in any sense.

The aim of this paper is to challenge this assumption. Our thesis is that slurs are always potentially harmful, even if some of their occurrences are nonderogatory. This does not mean that we take ourselves to be *prohibitionists* (see Anderson and Lepore 2013a, 2013b; see also Cepollaro, Sulpizio and Bianchi 2019, 33). That is, we do not think that it should be morally forbidden to utter a slur even when it is mentioned, for instance, for

¹ Other nonderogatory uses of slurs that have been discussed in the literature are referential (Anderson 2018) and identificatory (Zeman 2021) uses. We will briefly turn to these in section 3.

pedagogical purposes. But we do think that the utterance of a slur always comes at a moral cost, which in cases like this may be worth paying.

Our argument is the following. Derogatory occurrences of slurs (which, following Hom (2010), we call “orthodox occurrences”) are not characterized by their sharing any specific linguistic form; instead, they are those that take place in what we call *uncontrolled contexts*, that is, contexts in which we do not have enough knowledge of our audience to predict what the uptake of the utterance will be. Slurs uttered in *controlled* contexts, by contrast, may lack derogatory character. However, although the kind of context at which the utterance of a slur takes place can make it nonderogatory, it cannot completely deprive it of its harmful potential. Utterances of slurs in controlled contexts still contribute to *normalizing* their utterances in uncontrolled contexts, which makes nonorthodox occurrences of slurs potentially harmful too. It is not one of the aims of this paper to establish what makes utterances of slurs in uncontrolled contexts derogatory, nor in what sense exactly they are harmful. We just assume the common intuition that most occurrences of slurs are derogatory and in consequence harmful, and suggest that these coincide with those that take place in uncontrolled contexts. Our argument should be read as the conditional one that, *if* these occurrences of slurs are harmful, *then* all of them are potentially so.

Insofar as one of the outcomes of our work concerns the moral permissibility of mentioning a slur, it points in the same direction as Herbert (ms.), who argues that we should be careful even when merely *talking about* slurs. Her argument is that, just by mentioning these words, we already trigger harmful implicit associations.² Although we share Herbert’s concerns and reach a conclusion similar to hers, there are some differences between her work and ours that are worth commenting on. We will do so after presenting our argument.

The structure of the paper is as follows. In section 2, we review the distinction between derogatory and nonderogatory occurrences of slurs in terms of Hom’s (2010) distinction between orthodox and nonorthodox occurrences of slurs. We argue that the difference between these two kinds of occurrences does not lie in the linguistic form of the sentence uttered, but

² We are greatly indebted to Cassie Herbert for kindly sharing her manuscript with us.

in the context at which each of them takes place. In section 3, we flesh out what exactly this supposes by distinguishing between controlled and uncontrolled contexts; occurrences of slurs in uncontrolled contexts are always derogatory, while occurrences in controlled contexts may not be so. In section 4, we argue that occurrences of slurs in controlled contexts, even if sometimes not derogatory, have a normalizing potential that makes occurrences of slurs in uncontrolled contexts more likely. Hence, that an occurrence of a slur is nonderogatory does not mean that it does not have the potential to harm. We end the section by comparing our view with Herbert's (ms.). In section 5, finally, we discuss some of the consequences that our point may have for philosophical practice.

2. Nonorthodox occurrences of slurs

In this section, we survey the different cases in which occurrences of slurs have been said to be nonderogatory. These will be the cases on which we will focus in subsequent sections to discuss whether the fact that they are not derogatory means that they are not problematic in any sense. At the end of the section, we will argue that these cases include *uses* of slurs, despite attempts to reduce all nonderogatory occurrences of slurs to cases of *mention*.

According to Hom (2010, 168–69), some occurrences of slurs (which he calls “orthodox occurrences”) are nondisplaceable, while others (which he calls “nonorthodox occurrences”) are displaceable. Orthodox occurrences are nondisplaceable because they are derogatory even when embedded, while nonorthodox occurrences are displaceable in the sense that they are not always derogatory. In this paper, we do not embrace a particularly precise conception of derogation. It should be enough to say that derogation is the application to an individual of a negative moral evaluation (Hom and May 2013, 310), which is “an objective feature of the semantic contents of pejorative terms”³ (Hom 2012, 397). This distinguishes derogation from

³ This is not incompatible with accepting that some occurrences of slurs are derogatory and others are not. An occurrence of a slur can be nonderogatory because its

mere offense, which is a psychological phenomenon depending on the beliefs and values of participants in the conversation. Moreover, slurs, are opposed to mere insults, derogate its target in virtue of their belonging to a certain social group. We will assume that this kind of derogation is harmful in some way—be it because it subordinates (Kukla 2018) or dehumanizes (Jeshion 2018) its target. Let ‘*S*’ be a slur, and let us substitute it for the word that Hom uses in his examples of orthodox occurrences of slurs:

- (1) If there are *Ss* in the building, then *X* will be relieved.
- (2) There are no *Ss* in the building.
- (3) Are there *Ss* in the building?
- (4) John said that there are *Ss* in the building.
- (5) John said: ‘There are *Ss* in the building.’
- (6) In the novel, there are *Ss* in the building.

Hom takes ‘*S*’ to be derogatory in (1–6).⁴ He takes it to be so even if it appears in the antecedent of a conditional in (1), embedded under negation in (2), as part of a question in (3), reported in indirect style in (4) and in direct style in (5), and embedded under an “in the fiction” operator in (6) (see also Hornsby 2001, 129–130; Potts 2007, 166; Croom 2011, 347; Hom 2012, 384–385; Anderson and Lepore 2013a, 30; Croom 2014, 228). However, he does not take ‘*S*’ to be derogatory in his example of a nonorthodox occurrence of a slur, which he takes from (Hom 2008, 429) and we reproduce here substituting ‘*N*’ for the (alleged) neutral counterpart of ‘*S*’:⁵

- (7) Institutions that treat *Ns* as *Ss* are morally depraved.

semantic content does not result in derogation when it interacts with features of the particular linguistic environment or context of utterance.

⁴ As Hom (2010, n. 17) acknowledges, occurrences of ‘*S*’ such as the one in (5) are not incontrovertibly derogatory. To support the claim that they are, Hom argues that a speaker who is not a member of the target group and is not racist would be reluctant to utter (5) in front of a member of the target group, and that this is at least partly explained by the fact that the occurrence of the slur in it is derogatory. This argument seems sound enough to us.

⁵ We do not entirely agree with the idea that slurs have neutral counterparts, but we will assume throughout this paper that they do. For discussion on this issue, see Mühlebach (2019).

Hom takes (7) to contain a nonderogatory occurrence of ‘*S*’. Note that, in this case, ‘*S*’ is not embedded in any of the ways depicted in (1–6). In this paper, we will assume that Hom’s classification coincides with the intuitions of most speakers. Thus, we will take Hom’s diagnosis that (1–6) contain derogatory occurrences of ‘*S*’ and (7) does not as part of our data.

A couple of categories can be distinguished within nonorthodox occurrences of slurs. One of the cases that most readily come to mind is that of appropriated uses of slurs. Appropriated uses of slurs are those that take place when speakers belonging to the target group aim at demarcating the group or fostering solidarity or feelings of belonging (see Bianchi 2014; Cepollaro 2017; Anderson 2018). Speakers who have appropriated a slur can use it to refer to themselves or other members of the group without derogating anyone.

But we can also find occurrences of slurs that are nonorthodox without being instances of appropriation. These have been aptly labeled “nonderogatory, nonappropriated” (NDNA) uses of slurs, an umbrella term for all nonorthodox uses of slurs that are not appropriated (Hom 2008; see also Croom 2011). An example of an NDNA use of a slur (given by Hom 2008, 429) is:

(8) There are lots of *Ns* at *Y*, but no *Ss*.

(7) would be another example of an NDNA use of a slur.

Appropriated and NDNA uses are both *uses* of slurs, but paradigmatic nonorthodox occurrences of slurs are *mentions* of them. In fact, Hornsby apparently endorses the idea that all nonorthodox occurrences of slurs are at the end of the day cases of mention:

Certainly there are occurrences of derogatory words that are utterly inoffensive: “He is not [an *S*]” can be said in order to reject the derogatory “[*S*]”; one can convey that “[*S*]” is not something one calls anyone by saying “There aren’t any [*Ss*].” But these examples do not count against their uselessness as I mean this, because they are examples in which it is part of the speaker’s message that she has no use for the word “[*S*]”. We might gloss the two sentences so that the word is mentioned rather than used: “[“[*S*]” is not what he ought to be called]; “[“[*S*]” has no application.” (Hornsby 2001, 129)

Hornsby seems to reduce all cases of nonorthodox occurrences of slurs to cases of mention rather than use. A plausible paraphrase of (7) in which ‘*S*’ is mentioned rather than used would be this:

- (9) Institutions that treat *Ns* as deserving to be called ‘*Ss*’ are morally depraved.

However, we do not think that Hom’s distinction between orthodox and nonorthodox occurrences of slurs should coincide with the distinction between use and mention, so that every sentence featuring a nonorthodox occurrence of a slur can be paraphrased as a case of mention. In fact, we do not think that the distinction between orthodox and nonorthodox occurrences of slurs is a *linguistic* distinction, in the sense that a criterion to distinguish the latter from the former can be given just in terms of the form of the sentence used. A sentence in which a slur is used can be derogatory or nonderogatory independently of whether the slur appears in the antecedent of a conditional, embedded under negation, or as part of a question, and the same happens when the slur is merely mentioned.

The relevant factor when distinguishing between orthodox and nonorthodox occurrences of slurs is the context. (1) and (7) are both cases of use, but (1) is derogatory and (7) is not. (5) and (9) are both cases of mention, but (5) is derogatory and (9) is not. The differences lie in the kind of context that we most plausibly associate with each sentence: (1) is most easily imagined as uttered in a context in which an act of derogation takes place, while (7) tends to make us picture a context in which the speaker is in fact denouncing derogatory practices, and something parallel to this can be said about (5) and (9).

Thus, the difference between orthodox and nonorthodox occurrences of slurs does not lie in the form of the sentence used, but in the context in which they take place. In the next section, we flesh out what exactly distinguishes some contexts from others. However, as we will see, the difference has a limited impact, since it makes nonorthodox occurrences of slurs less dangerous, but not strictly not dangerous.

3. Controlled and uncontrolled contexts

The upshot of the previous section was that whether an occurrence of a slur is orthodox or nonorthodox depends on the context at which it takes place. In this section, we take a closer look at the kinds of contexts that make an occurrence of a slur orthodox or nonorthodox. In particular, we identify nonorthodox occurrences of slurs with those that can take place in what we call “controlled contexts” and orthodox occurrences of slurs with those that take place in “uncontrolled contexts”. However, we will see in section 4 that, even if the distinction between controlled and uncontrolled contexts can help us rank occurrences of slurs according to their derogatory character, part of slurs’ power to cause harm is distributed equally across the categories distinguished here.

Communication is a risky business. There are a number of factors that can have an impact on the kind of effect that a given utterance will have, and most of them escape our control. When communicating, we often have to manage without knowing what our audience knows or what their expectations are. Still, even if rare, contexts can be found in which we can predict with reasonable accuracy what the consequences of a given utterance will be. We call these “controlled contexts”.⁶ When we are talking about utterances including a slur, an example of a controlled context would ideally be a pedagogic one, and another, more contentious one would be that in which a slur is successfully used in an ironic way.

Here is an example of a pedagogic occurrence of a slur. Our son Dani comes home from school and says his friend *Y* says his other friend *X* is an *S*. Later on, we tell Dani he should never say that word again. ‘What word?’, he says. He has not forgotten it, but honestly cannot recall which one of the words he has pronounced we are forbidding him from saying. We feel forced to pronounce ‘*S*’ in order to make sure he knows what term we are referring to, so we do—we say ‘We don’t call people ‘*S*’, that’s an ugly thing to say.’ We have uttered a slur, even if we have only mentioned

⁶ Of course, whether a given context is a controlled one will in many cases not be a settled matter. We leave for further work to offer precise criteria for a context to fall under this label.

it.⁷ But we had no other option, and we can be sure that by doing this we have not insulted anyone—if anything, we have prevented Dani from insulting anyone, even if from unintentionally doing so. We know enough about our own son to guarantee that he has understood that we were not insulting anyone. Here, the occurrence of ‘*S*’ is nonderogatory.

Here is another kind of case in which we can say that the utterance of a slur has taken place in a controlled context. This time, we are not talking about a mere mention of a slur, but about a full-blown use—an ironic use. We are a progressive group of friends who would never as much as mention a slur in front of strangers, much less use it to insult a person on grounds of her belonging to a given group. However, we find fun in imitating bigots’ mannerisms, and enjoy inner jokes that include ironic uses of ‘*S*’. We are completely sure that all our friends in the group share our sensibility, and that none of them will take us to aim at insulting anyone. We think it is intuitive to take occurrences of slurs such as these to be nonderogatory, whatever the form of the sentences in which they appear.

Other nonderogatory uses of slurs that have recently been described are referential (Anderson 2018) and identificatory (Zeman 2021) uses. Referential uses take place when members of the target group use a slur to address other members without any intention to appropriate the term, while identificatory uses take place when they simply take the word to be the one that refers to the group they take themselves to belong to. We take these uses to take place in controlled contexts too, as the speaker’s group membership is salient enough for her to be confident that the audience will understand that she did not mean to insult, just like happened in Dani’s case.

If controlled contexts are those in which we can be sure about the other participants’ knowledge and expectations, almost all contexts in which we can find ourselves are uncontrolled ones. It is difficult to know anyone as

⁷ It could be argued that pedagogic contexts not only allow for nonderogatory *mentions* of slurs, but also for nonderogatory *uses* of them. For instance, someone might claim, we could have also said to Dani ‘There are no *S*s, only *N*s.’. However, we find it hard to see this sentence as nonderogatory—the fact that the speaker has seen the need to categorize Dani’s friend as belonging to the target group, even if in a supposedly neutral way, makes the sentence problematic. See again Mühlebach (2019).

well as we know our own children or our closest friends, and in many cases we hardly have any relevant information about our audience. Consider our daily interactions with strangers, and the limiting case of the completely uncontrolled context in which public communication takes place. When we utter a slur in an uncontrolled context, our audience has every reason to attribute to us a negative attitude toward a given group, and we cannot reasonably expect not to be attributed such an attitude, which is what, in an intuitive sense, means to derogate (see section 2). Thus, in uncontrolled contexts, which are most of the contexts, occurrences of slurs are derogatory.

As advanced before, it lies beyond the scope of this paper to offer an explanation of exactly how slurs derogate when uttered in an uncontrolled context. There are a number of proposals in the market that aim at accounting for this fact. Some of these views rely on specifically derogatory content that is part of what is said (Hom 2008, 2010, 2012; Hom and May 2013), conventionally implicated (Potts 2007; Copp 2009; Williamson 2009; McCready 2010; Whiting 2013), or presupposed (Macià 2002; Schlenker 2007; Cepollaro and Stojanovic 2016; Marques and García-Carpintero 2020). Some of them, by contrast, explain the derogatory character of slurs without appealing to specifically derogatory content (Anderson and Lepore 2013a, 2013b). At any rate, these are all different ways of accounting for the widely held intuition that, in most of the cases, occurrences of slurs are derogatory.

A plausible objection is that slurs can occur in a derogatory way in controlled contexts too. I may know exactly what the reaction of the audience to my utterance of a slur will be, and know this reaction to be one that will precisely result in derogation. In this case, the occurrence of the slur will be derogatory even if it takes place in a controlled context. Note, however, that what distinguishes controlled and uncontrolled contexts is that occurrences of slurs in the former *can* be nonderogatory, not that they will always be so.

Another plausible objection, mirroring the one above, is the following. In uncontrolled contexts, we cannot be sure that the uptake of our utterance will fail to derogate, but this does not mean that it will derogate. It may happen that, just by chance, every single member of the audience

understands the occurrence of the slur as nonderogatory, even if we are not able to predict that this will be the case. For instance, all bystanders who hear Dani utter a slur could assume that he does not know what the word means.⁸ Note, however, that we have characterized uncontrolled contexts as those in which hearers have *every reason* to attribute to us a negative attitude toward the target group, which we cannot in turn *reasonably* expect not to be attributed to us. In this case, the audience can refuse to attribute the negative attitude to us. However, inasmuch as they would be warranted in so doing, our utterance can be taken to be derogatory.

Of course, what counts as good reason is a highly context-dependent issue, and some contexts might make it reasonable not to attribute a negative attitude to the speaker. This may be the case, for instance, when the audience knows that the speaker is a decent person. Note, however, that this will only be warranted if the audience knows not only that the speaker is a decent person, but also that the speaker is aware that this is publicly known. In this case, the audience will have reason to believe that the utterance is not derogatory, but we will no longer be facing an uncontrolled context. If the relevant piece of public knowledge is missing, as should happen in an uncontrolled context, and the speaker still chooses to utter the slur, the audience can legitimately conclude that she is comfortable with being attributed a negative attitude.

This idea that, in uncontrolled contexts, it is reasonable to attribute the utterer of a slur a negative attitude toward the target group no matter what her actual attitudes are is similar to one that has been defended by Lasersohn (2007). This idea is a key component in Lasersohn's explanation of the hyperprojectivity of slurs' derogatory content. Hyperprojectivity is the phenomenon whereby the derogatory content of slurs is able, in many occasions, to survive in grammatical constructions that would usually block presuppositional content. Lasersohn defends that this fact is compatible with a presuppositional account of slurs by providing the following explanation. According to Lasersohn, slurs are emotionally charged terms, so uttering them entails a social risk. Lasersohn believes that speakers are aware of the social burden of slurs, and this is the main reason why most

⁸ Thanks to an anonymous reviewer for *Organon F* for suggesting this objection to us.

speakers avoid uttering slurs—because, whatever their particular attitudes, they are aware that they can reasonably be attributed bigotry. Precisely because of this, when a speaker does utter a slur, it makes sense to think that she is comfortable with being identified as a bigot, and this is how the derogatory content of the slur projects where most presupposed content does not (Lasersohn 2007, 228). Like Lasersohn, we think that, if a speaker utters a slur in an uncontrolled context in which it is even merely possible that someone understands the occurrence of the slur as derogatory, it makes sense to take the speaker to be comfortable with this possibility, and thus to take the occurrence to be actually derogatory.

4. The normalizing potential of slurs

We have seen that slurs are derogatory in uncontrolled contexts, but not in controlled contexts such as pedagogic and ironic ones. Still, no matter how carefully we arrange the current context to make sure that the utterance of a slur does not have the kind of effect we want to avoid, it will facilitate ulterior occurrences of the term. In particular, it will make the slur more likely to appear in uncontrolled contexts in which the utterance of the slur is derogatory. In this section, we explore how this could be the case with the two kinds of occurrences of slurs that we presented in the previous section—pedagogic mentions and ironic uses of slurs. If even the apparently most “innocent” occurrences of slurs, such as those that take place in pedagogic contexts, are potentially harmful, it is natural to conclude that *all* occurrences of slurs are potentially harmful.

Let us start with the irony case. Remember that, in this case, we are a progressive group of friends who enjoy using slurs in an ironic way to make fun of bigots. However, we have ironically used ‘*S*’ in our friend group so many times that we have deprived it of its forbidden character—it no longer makes us uncomfortable to hear the word, which makes its utterance in uncontrolled contexts more likely now. This nonderogatory use of a slur thus normalizes derogatory occurrences of the word, and is potentially harmful in this sense. Of course, our friend group could be careful enough not to let uses of slurs slip out of the controlled context. This is why ironic uses of slurs such as these are not harmful *tout court*, but potentially

harmful. But, since potential harm implies actual danger, these uses are dangerous *tout court*.

Now, take the example involving a pedagogic mention of a slur, also described in the previous section. Remember that, in this case, we feel forced to utter the word ‘*S*’ in order to make our son Dani aware that he should not call anyone an *S*. However, we have taught Dani what ‘*S*’ means, thus giving him the tools to use the word to insult if he wants to do so at some point. Note that, at least in this case, the risk that Dani grows up to use ‘*S*’ as an insult may be worth it: as we will see, preventing an actual risk may be preferable to preventing a virtual one. In this sense, this case might strike us as clearer than the previous one. The normalizing potential is similar in both cases, though. The difference is that, in the pedagogic case, it is clearer how the benefits could outweigh the potential harm. In the irony case, all we have on the positive side is the fun we have with our friends. Referential and identificatory uses of slurs are closer to pedagogic mentions than to ironic uses in this respect. Like with pedagogic mentions, however, there is still the risk that these uses facilitate ulterior occurrences of the term in contexts in which the group membership of the speaker, although salient, does not make the audience understand such occurrences as nonderogatory because the speaker does not belong to the target group. Hence, the moral here is that we may have full control over the present context, but we do not have control over all possible future contexts. Thus, slurs always have normalizing potential. The slur might not be problematic in the context at which it is uttered, but it may reveal itself to be so as we look beyond the original conversation and consider other exchanges that might be facilitated by the original utterance. We take something to be dangerous whenever it *may* cause harm, even if it does not actually do so. Insofar as occurrences of slurs are always potentially harmful, therefore, we take them to be always dangerous.

As we said in the introduction to this paper, our idea that even non-derogatory occurrences of slurs can be harmful should not suffice to classify us as *prohibitionists* (see Anderson and Lepore 2013a, 2013b; see also Cepol-
laro, Sulpizio and Bianchi 2019, 33)—we do not think utterances of a slur should be forbidden *tout court*. We think there are some practical consequences to the categorization of some uses of slurs as appropriated or

NDNA, and that the distinction between use and mention has practical consequences when it comes to slurs too. It may be permissible to mention a slur in certain contexts, just like it may be permissible to make an appropriated or an NDNA use of a slur. This marks a difference between these occurrences of slurs and their full-blown, derogatory uses. But we should be aware that these practices come with a moral cost too. The price may be worth paying, of course. It just misrepresents our moral life to assume that it consists in choosing the only permissible thing to do in each case; rather, we assess the moral costs and benefits of each course of action, decide what weight to give to each, and act in consequence.⁹ The moral benefits of performing a certain utterance of a slur might outweigh the pervasive moral cost we have described, and so it might be worth it to utter the slur.

As we also advanced in the introduction to this paper, the stance that merely mentioning a slur, as we do in the pedagogic case, can be reprehensible too has been defended before us. Herbert (ms.) argues that the practice of offering examples of slurs, which is widespread in philosophy, may cause harm just like using them does. To conclude this, she relies on empirical evidence found by Carnaghi and Maass (2008) and Fasoli, Paladino, Carnaghi, Jetten, Bastian and Bain (2015) that it is mere exposure to a slur, rather than specifically exposure to *uses* of a slur, that triggers negative implicit associations concerning the target group. The question that Herbert asks herself (is it morally permissible to mention a slur?) is precisely one of those we have set ourselves to answer in this paper, and the reply she offers is akin to ours—a refusal to give a context-independent answer, together with an invitation to be extremely careful when deciding whether to mention a slur. However, there are some differences between Herbert’s work and ours that we think make this paper a worthy contribution to the debate.

First, Herbert’s work focuses on mentions of slurs, while ours also covers some of their uses, such as ironic ones. Of course, this does not mean that Herbert’s point applies only to mention. The way her argument goes, mentions of slurs are potentially harmful just in virtue of their being *occurrences* of slurs, so her conclusion should apply to any occurrence of these words, including uses in general and ironic uses in particular. But, while most of

⁹ For a really insightful guide on what particular factors to consider when deciding whether a slur is worth mentioning, see Herbert (ms.).

the discussion in Herbert's paper concerns mentions of slurs in academic and journalistic environments, ours concerns ironic uses and mentions of slurs in *pedagogic* contexts.

Second, our explanation of how these occurrences of slurs can end up being harmful is more general than Herbert's. While she relies on implicit associations, our argument is compatible with different proposals as to what mechanism accounts for the pernicious effects that uttering a slur may have, possibly including the appearance of implicit associations like the ones described by Herbert. If (contrarily to the evidence we now have) occurrences of slurs turned out not to elicit implicit associations *per se*, our work would still provide a schema that could be completed with an alternative mechanism.

5. Concluding remarks

In this paper, we have argued that the mere utterance of a slur has a certain kind of impact—it normalizes further occurrences of the word. This is so even in those cases in which the slur does not derogate anyone. Nothing prevents the slur from being used in a derogatory way in the ulterior occurrences normalized by these ones, so even nonderogatory occurrences make it more likely for derogation to take place at some point in the future.

Our proposal has consequences for philosophical practice as we know it. Our point is that mentions of slurs in academic papers are potentially harmful too: even if, not being used, they are not derogatory, they facilitate ulterior occurrences of the slur in question too. Of course, not all philosophers are comfortable with mentioning slurs even if they take it to be necessary. Rebecca Kukla, for instance, says:

By flagging that I will be mentioning slurs and reminding the reader that even the mention of slurs can harm, I hope to frame these mentions in a way that allows readers to be conscious of such effects and to try to minimize it. I also use scare quotes around the slurs throughout, to help avoid normalizing them as part of everyday speech, and in the hope of marking them at the

visual level as problematic terms that I am not uttering in my own voice and that are not to be taken for granted as readable. (Kukla 2018, 24)

We do not think the use of scare quotes blocks normalization, but again, this does not mean that Kukla's mentions of slurs are necessarily unjustified. The moral cost might be worth assuming in this case. We keep our doubts, however, that it is worth assuming in cases in which more examples than the strictly necessary are given.

We do not think that a criterion can be found by which certain occurrences of a slur should be allowed, but by the same token we do not think that there is a criterion that forbids the rest of occurrences *tout court*. Our point is that uttering a slur always comes at a moral cost, and it is the responsibility of the speaker, or the philosopher who writes a paper on slurs, to assess such cost and decide whether it is worth it to mention a word to explain to a child that it should never be used or to give one more example of a slur in a paper addressed to an audience that is assumed to know what slurs are.

Acknowledgments

This paper has been funded by the Spanish Ministry of Science and Innovation under the research project "Disagreement in Attitudes: Normativity, Affective Polarization and Disagreement" (PID2019-109764RB-I00), by the Regional Government of Andalusia under the research projects "Public Disagreements, Affective Polarization and Immigration in Andalusia" (B-HUM-459-UGR18) and "The Inferential Identification of Propositions: A Reconsideration of Classical Dichotomies in Metaphysics, Semantics and Pragmatics" (P18-FR-2907), and by the University of Granada under a "Contrato Puente" fellowship and the excellence unit FiloLab-UGR (UCE. PPP2017.04). The authors would also like to thank Alex Davies, María José Frápolli, Andrés Soria, Neftalí Villanueva, Dan Zeman, and two anonymous reviewers for *Organon F*, as well as audiences at EvalLang-2019, FiloLab International Summer School 2019, Epistemological and Cognitive Analyses of Cognition, Beliefs and Knowledge, and the IX Meeting of the Spanish Society for Analytic Philosophy, for their helpful comments and suggestions.

References

- Anderson, Luvell. 2018. "Calling, Addressing, and Appropriation." In *Bad Words: Philosophical Perspectives on Slurs*, edited by David Sosa, 6–28. Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780198758655.003.0002>
- Anderson, Luvell, and Ernie Lepore. 2013a. "Slurring Words." *Noûs* 47 (1): 25–48. <https://doi.org/10.1111/j.1468-0068.2010.00820.x>
- Anderson, Luvell. 2013b. "What Did You Call Me? Slurs as Prohibited Words." *Analytic Philosophy* 54 (3): 350–63. <https://doi.org/10.1111/phib.12023>
- Bianchi, Claudia. 2014. "Slurs and Appropriation: An Echoic Account." *Journal of Pragmatics* (66): 35–44. <https://doi.org/10.1016/j.pragma.2014.02.009>
- Carnaghi, Andrea, and Anne Maass. 2008. "Derogatory Language in Intergroup Context: Are "Gay" and "Fag" Synonymous?" In *Stereotype Dynamics: Language-Based Approaches to the Formation, Maintenance, and Transformation of Stereotypes*, edited by Yoshihisa Kashima, Klaus Fiedler, and Peter Freytag, 117–34. New York: Lawrence Erlbaum Associates.
- Cepollaro, Bianca. 2017. "Slurs as the Shortcut of Discrimination." *Rivista di Estetica* (64): 53–65. <https://doi.org/10.4000/estetica.2063>
- Cepollaro, Bianca, and Isidora Stojanovic. 2016. "Hybrid Evaluatives: In Defense of a Presuppositional Account." *Grazer Philosophische Studien* 93 (3): 458–88. <https://doi.org/10.1163/18756735-09303007>
- Cepollaro, Bianca, Simone Sulpizio, and Claudia Bianchi. 2019. "How Bad Is It to Report a Slur? An Empirical Investigation." *Journal of Pragmatics* (146): 32–42. <https://doi.org/10.1016/j.pragma.2019.03.012>
- Copp, David. 2009. "Realist-Expressivism and Conventional Implicature." *Oxford Studies in Metaethics* 4: 167–202. <https://doi.org/10.1017/S0265052500002880>
- Croom, Adam M. 2011. "Slurs." *Language Sciences* 33 (3): 343–58. <https://doi.org/10.1016/j.langsci.2010.11.005>
- Fasoli, Fabio, Maria Paola Paladino, Andrea Carnaghi, Jolanda Jetten, Brock Bastian, and Paul G. Bain. 2015. "Not "Just Words": Exposure to Homophobic Epithets Leads to Dehumanizing and Physical Distancing from Gay Men." *European Journal of Social Psychology* 46 (2): 237–48. <https://doi.org/10.1002/ejsp.2148>
- Herbert, Cassie. ms. "Talking About Slurs."
- Hom, Christopher. 2008. "The Semantics of Racial Epithets." *The Journal of Philosophy* 105 (8): 416–40. <https://doi.org/10.2307/20620116>
- Hom, Christopher. 2010. "Pejoratives." *Philosophy Compass* 5 (2): 164–85. <https://doi.org/https://doi.org/10.1111/j.1747-9991.2009.00274.x>
- Hom, Christopher. 2012. "A Puzzle About Pejoratives." *Philosophical Studies* 159 (3): 383–405. <https://doi.org/10.1007/s11098-011-9749-7>

- Hom, Christopher, and Robert May. 2013. "Moral and Semantic Innocence." *Analytic Philosophy* 54 (3): 293–313.
<https://doi.org/https://doi.org/10.1111/phib.12020>
- Hornsby, Jennifer. 2001. "Meaning and Uselessness: How to Think About Derogatory Words." *Midwest Studies in Philosophy* 25 (1): 128–41.
<https://doi.org/10.1111/1475-4975.00042>
- Kukla, Rebecca. 2018. "Slurs, Interpellation, and Ideology." *The Southern Journal of Philosophy* 56 (S1): 7–32. <https://onlinelibrary.wiley.com/doi/10.1111/sjp.12298>
- Lasersohn, Peter. 2007. "Expressives, Perspective, and Presupposition." *Theoretical Linguistics* 33 (2): 223–30. <https://doi.org/10.1515/TL.2007.015>
- Macià, Josep. 2002. "Presuposición y significado expresivo." *Theoria: Revista de Teoría, Historia y Fundamentos de la Ciencia* 3 (45): 499–513.
- Marques, Teresa, and Manuel García-Carpintero. 2020. "Really Expressive Presuppositions and How to Block Them." *Grazer Philosophische Studien* 97 (1): 138–58. <https://doi.org/10.1163/18756735-09701008>
- McCready, Elin. 2010. "Varieties of Conventional Implicature." *Semantics and Pragmatics* 3 (8): 1–57. <https://doi.org/10.3765/sp.3.8>
- Mühlebach, Deborah. 2019. "Semantic Contestations and the Meaning of Politically Significant Terms." *Inquiry*.
<https://doi.org/10.1080/0020174X.2019.1592702>
- Potts, Christopher. 2007. "The Expressive Dimension." *Theoretical Linguistics* 33 (2): 165–98. <https://doi.org/10.1515/TL.2007.011>
- Schlenker, Philippe. 2007. "Expressive Presuppositions." *Theoretical Linguistics* 33 (2): 237–45. <https://doi.org/10.1515/TL.2007.017>
- Whiting, Daniel. 2013. "It's Not What You Said, It's the Way You Said It: Slurs and Conventional Implicatures." *Analytic Philosophy* 54 (3): 364–77.
<https://doi.org/10.1111/phib.12024>
- Williamson, Timothy. 2009. "Reference, Inference, and the Semantics of Pejoratives." In *The Philosophy of David Kaplan*, edited by Joseph Almog, and Paolo Leonardi, 137–59. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195367881.003.0009>
- Zeman, Dan. 2021. "A Rich-Lexicon Theory of Slurs and Their Uses." *Inquiry*.
<https://doi.org/10.1080/0020174X.2021.1903552>