



How slurs enact norms, and how to retract them

Teresa Marques¹

Received: 20 April 2022 / Accepted: 3 April 2024
© The Author(s) 2024

Abstract

The present paper considers controversial utterances that were erroneously taken as derogatory. These examples are puzzling because, despite the audiences' error, many speakers retract and even apologise for what they didn't say and didn't do. In recent years, intuitions about retractions have been used to test semantic theories. The cases discussed here test the predictive power of theories of derogatory language and help us to better understand what is required to retract a slur. The paper seeks to answer three questions: are the cases considered genuine retractions? If the speakers didn't derogate by using a slur, how are the cases retractions of *derogatory acts*? Do these examples support expressive accounts of slurs? I argue that the examples provide evidence for an expressivist account of slurs: a slurring utterance (defeasibly) makes a derogatory speech act where the speaker expresses a commitment to a morally questionable appraisal state, such as disgust or contempt for a target group. A retraction of a derogatory speech act requires undoing the enactment of that commitment, which can be achieved with a genuine apology. This helps explain the conduct of audiences who misunderstand what the speaker says, and the speakers' reactions.

Keywords Expressives · Slurs · Derogation · Apology · Retraction

1 Contentious cases: much ado about... what?

A high school geography class about African countries, a lecture on business communication in China, a discussion about a budget with Washington DC's mayor, a Latin-American footballer's expression of gratitude towards a fan, a K-Pop band's new love song. This is not the start of a fictional work of magical realism. It is a list of real cases where speakers were erroneously accused of *using* a racial slur. The cases are real, but the public reactions to such mundane events, and subsequent developments, are somewhat surreal. Those accused were, in many of the cases, required to retract,

✉ Teresa Marques
teresamatosferreira@ub.edu

¹ Philosophy Department, University of Barcelona, LOGOS Research Group in Analytic Philosophy, BIAP–Barcelona Institute of Analytic Philosophy, C/Montalegre, 6–8, 08001 Barcelona, Spain

apologize, or otherwise punished. In this paper, I argue that one can make sense of these controversial examples by adopting an expressivist account of slurs: that a slurring utterance (defeasibly) makes a derogatory speech act whereby the speaker expresses a commitment to a morally questionable appraisal state, such as disgust or contempt for a target group. A retraction of a derogatory speech act requires undoing the enactment of that normative appraisal, something that can be achieved with a genuine apology. This can help explain the conduct of audiences, and speakers' responses.

This brief introduction raises three questions, which the paper seeks to answer:

- (i) Are all cases considered genuine instances of retraction?
- (ii) If the cases considered are not cases of slurs, in what sense (if any) is the paper still about the retraction of slurs?
- (iii) Do the cases considered support expressive accounts of slurs?

Question (ii) can receive a preliminary reply here. The examples I discuss parallel classic cases of misperception and error. Cases of error have been central in philosophy. In his "Morally Loaded Cases in Epistemology", for instance, Williamson (2019) addresses arguments about *seemings*, where internalists about epistemic justification would say:

Whether a belief is justified at a time depends on its coherence with the internal consciously introspectable mental states of the subject at that time, especially *seemings*, and perhaps other beliefs too. Seemings are pre-doxastic; they are neither beliefs nor inclinations to believe. You have a seeming when things seem to you a certain way, either sensorily or intellectually. Seemings can be false: sometimes things seem to you to be some way even though they are not in fact that way. Still, when it seems to you that *P*, you are at least *prima facie* justified in believing that *P*. You are all-things-considered justified in believing that *P* when so believing also coheres with your other relevant mental states, especially your seemings. Consequently, false beliefs are sometimes justified. (Williamson, 2019, p. 4)

Externalists like Williamson deny this conclusion. But even if a false seeming does not give *prima facie* justifying reason for belief, it may nonetheless provide an *explanatory* reason for why the subject acts as he does. The content of the *seeming* must be thus considered in the rational explanation of the subject's conduct.

The real-life cases discussed here also involve *seemings* and are morally loaded. The errors concern the seeming performance of certain speech acts. Audiences misidentify the words used as slurs, misinterpret the meaning of those words, mistake the illocutionary act performed, and misjudge the speaker's intentions. Despite their error, audiences' subsequent behaviour should be rationally explained.

In the rest of this section, I introduce the cases that concern us here. In the second section, I make a non-exhaustive introduction to theories of slurs, contrasting some semantic and pragmatic accounts. In the third section, I argue that an Austinian understanding of retraction as the undoing of the normative commitments enacted by a previous speech act (after Caponetto, 2020) serves well to explain the retraction of derogatory expressive speech acts. In particular, the retraction of *derogatory expressive speech acts* can be expected to differ from that of *constative acts*. This will help

in answering question (i): not all the cases considered count as intended retractions of derogations, although some are.¹ However, as I'll argue, a genuine apology should be considered as an indirect retraction of the morally problematic act of derogation. Finally, to answer (iii): expressive theories of slurs (as expressive of derogation) are better positioned than others to rationally explain the conduct of audiences that demand that speakers retract, whether the speaker uttered a derogatory utterance or whether the audience is mistaken in so thinking. Had the speaker merely communicated an at-issue (or not-at-issue) proposition p , its retraction would not require the manifest admission of moral blame for the wrong of derogation that is expressed through a genuine apology. And if the audience had merely misunderstood the speaker as having communicated an at-issue (or not-at-issue) proposition, then the audience *ought* to have accepted that they just misunderstood what had been said. But these audiences do not accept such corrections. Something else must be going on.

1.1 Mentioning slurs

An article in *The Times* from June 2021 informs that a high-school teacher, Mr. Collins, tried to show his students how to pronounce 'Niger'. Collins pointed out that the country's name rhymes with 'Pierre', and that its pronunciation should not be confused with that of the racial epithet, which he mentioned. The article reports that, according to a source, the teacher did not intend to be offensive and that "He was just trying to ensure nobody made a mistake that could be offensive." Note that the article, and the head teacher in a letter to parents, misleadingly said that the teacher *used* a slur, without making clear that it was a *quotational* use, thus suggesting that it was a standard derogatory one:

The teacher has been described as highly regarded and conscientious and he later apologised to the class. However, it was reported that teachers, parents and other pupils complained, with some apparently not realising the context in which the word had been used... During a lesson later the same day the children were handed complaint forms and told to write down comments if they were unhappy with Mr Collins's behaviour, a school source told the newspaper. In a letter to parents, Christian Kingsley, the head teacher, said that "a member of staff used racist language in a Year 9 lesson" and added that the matter would be investigated robustly. "The language was not directed at a student," the letter said. "However, we understand that students and their families have, justifiably,

¹ A reviewer disagrees that those cases count as retractions. In my view, the disagreement is partly terminological. A retraction is a speech act that targets a previously unretracted speech act. But if no previous speech act was performed (there was no actual derogation), then the speaker has nothing to take back. The reviewer's understanding rests on a restrictive interpretation of illocutionary felicity conditions: If the conditions fail, then the act is not performed. I think that this reading deprives us of the means to explain infelicitous acts: a promise to do what is not in one's power to fulfil, an order given by one lacking authority, an announcement that no one pays attention to, testimony that is a bald-face lie, or a confession of a crime that one is not responsible for. The rational explanation of people's conduct and of why they deserve criticism warrants seeing those cases as flawed speech acts. Hence, infelicitous retractions are also possible, as the cases here considered illustrate.

taken great offence to this and that there was no reason for the term being used.” (Woolcock, 2021)

One could think that this case supports the idea that the derogatory force of slurs always scopes out, even when they are mentioned.² If that were true, we would expect all mentions of slurs to be “uses of racist language”. That is disputable. A famous case of a mention of the n-word occurred during the *People v. O.J. Simpson* trial. Mark Fuhrmann, former detective of the Los Angeles Police Department and a crucial witness against Simpson, was cross-examined and accused of lying about his prior past use of the n-word. The defence attorney, F. Lee Bailey, mentioned the very word during questioning, and the cross-examination served to accuse Fuhrman of committing perjury. Recently, in the FX series *American Crime Story*’s episode about the trial, the n-word was fictionally used.³

If it were true that the derogatory force of slurs always scopes out, and if this were a reason to retract any mention of a slur, then many discussions of the meanings of slurs (including most of the papers cited here), in court cases (such as the O.J. Simpson trial), as well as in fiction (e.g., the series about the trial) should be retracted. This would make work in philosophy and linguistics more difficult, and it would make essential aspects of court trials, documentaries, and reality-involving works of fiction impossible, depriving people of important pieces of social and historical knowledge, and of the understanding of human nature that such works can provide. Worst, it would confound the moral difference between (apparent) racists and people who, from all available evidence, are quite the opposite (e.g., the difference between the former LAPD detective, Mark Fuhrman, and the high-school teacher, Mr. Collins). The director of the school, Mr. Kingsley, went on to say, “[w]hile this language was not directed at a student, the term being used in any context has understandably caused great offence within our community. We *apologised* to students and families and are investigating through our staff disciplinary policy.” (My emphasis).

1.2 Different etymology, different meaning, same language

On January 15, 1999, David Howard, aide to then mayor of Washington, D.C., Anthony Williams, used ‘niggardly’ in reference to a budget. According to *the Webster English Dictionary*, the first recorded use of ‘niggard’ is from the fourteenth century, as a noun meaning a stingy person or a miser. Its use as a verb (‘niggard’ and ‘niggardly’) date from the sixteenth century. The word is of Scandinavian origin, akin to Old Norse ‘hnøgggr’. It has no common etymology or history with the racial epithet, which evolved from the Latin word ‘niger’. As reported in an article in the *Washington Post*, Howard’s use of ‘niggardly’ upset one of his black colleagues, who misinterpreted it as the racial slur and lodged a complaint. As a result, on January 25, Howard presented his resignation, which the mayor accepted.

After public pressure, the mayor offered Howard the chance to return to his position, but Howard refused to return to his previous post. He felt that he had “learned from the

² This feature of slurs will be discussed in Sect. 2.

³ For discussion of the difference between the trial and the series, see Robinson (2016).

situation” (Woodlee, 1999). At the time, the NAACP⁴ Chairman Julian Bond criticised the DC mayor, and said that people should not have to “censor their language to meet other people’s lack of understanding” while praising Howard’s reinstatement in another office.

1.3 Same etymology, different meaning, different language

Edinson Cavani is a Uruguayan and a professional footballer. He used to play with Manchester United. In a match between Manchester United and Southampton at the end of 2020, Cavani scored two goals, ensuring Man. United’s victory (3–2). A Spanish speaking fan (probably from Uruguay) congratulated Cavani on Instagram writing the Spanish sentence in (1) below.

- (1) Así te quiero, matador! [literally, ‘this is how I want you, you killer!, meaning roughly ‘you rule, mate!’].

Cavani thanked the fan in an Instagram story with the Spanish sentence in (2) followed by a handshake emoji.

- (2) Gracias, negrito.

The media storm that followed included news articles and headlines in the British and international press, a sample of which are reproduced below. Cavani subsequently retracted, apologised, and accepted to be punished:

- Edinson Cavani facing FA⁵ probe after sparking racism storm for calling a fan ‘negrito’ in Instagram post just hours after inspiring Manchester United’s comeback win at Southampton. (Kajumba, 2020)
- Edinson Cavani apologises for using racial term in Instagram post.
- Striker insists he meant no offence and is backed by club.
- Cavani could still be hit with FA ban despite deleting post. (Jackson, 2020)
- Man United striker Edinson Cavani is handed a three-match ban by the FA for ‘Gracias negrito’ Instagram post. (Wheeler, 2020)

The reaction from the Latin-American press and organisations was radically different. The Academia de las Letras de Uruguay issued an official statement vehemently denouncing the FA for the three-game ban and fine against Cavani as culturally and linguistically ignorant. The Academia further stated:

Cavani’s use of the word *negrito*... has this kind of affectionate tenor: given the context in which it was written, the person to whom it was addressed and the variety of Spanish that was used, the only sense that *negrito* can have—and particularly because of its diminutive nature—is the affective one. This is the meaning that appears in the *Dictionary of the Spanish of Uruguay*—edited in 2011 by the National Academy of Letters of Uruguay—under the term *negro*.

Unfortunately, the English Football Federation seems to ignore the above, and is wrong when it applies rule E3, which seeks to prohibit its players from

⁴ NAACP is the acronym of the National Association for the Advancement of Colored People in the USA.

⁵ The FA is the British Football Association.

making comments referring to the “race, ethnicity, colour or nationality” of a person and to hold players responsible for their accounts on social networks, a commendable issue in general but which is clearly not applicable at all to the case under consideration.

It would be equally inappropriate for this Academy to sanction some groups of Río de la Plata speakers who use in social networks the vocatives rey [king] and reina [queen] invoking a monarchical pretension that Río de la Plata abandoned centuries ago. In short: the English Football Federation has committed a grave injustice . . . and has shown ignorance and error regarding the uses of the language and in particular of Spanish, without taking note of all its complexities and contexts. (Academia de las Letras de Uruguay, 2020. My translation from the original)

In Spanish, ‘negro’ and ‘negrito’, or ‘negra’ and ‘negrita’ are not racial slurs. They mean the same as ‘black’ in English. But Cavani is Urugayan, and I’m informed by Argentinian and Uruguayan Spanish speakers that at least in the regions around the Río de la Plata, which include Buenos Aires and Montevideo, “negro” or “negrito” is not used co-extensively with “black person”. It is rather predominantly a term of endearment, something akin to “mate”, as the Uruguayan Academy also claims. It’s understandable that Uruguayans should be offended by Cavani’s punishment with a 3-game suspension and a £100.000 fine for something he didn’t say, didn’t mean and didn’t intend. So, why did he accept to retract and to apologise?

1.4 Different etymology, different meaning, different language

A similar controversy arose when a professor of Clinical Business Communication at the University of Southern California, Greg Patton, used a Chinese filler word during a zoom class. Patton was teaching a class in an online course about the use of “filler words” in speech. He said: “If you have a lot of ‘ums and errs,’ this is culturally specific, so based on your native language. Like in China, the common word is ‘that, that, that.’ So in China, it might be ‘nèi ge, nèi ge, nèi ge’.”⁶

In a letter sent to the school administration, Patton was accused of “using Chinese words that sounded like an English racial slur.” (Yeung, 2020):

“This phrase, clearly and precisely before instruction is always identified as a phonetic homonym and a racial derogatory term, and should be carefully used, especially in the context of speaking Chinese within the social context of the United States,” according to the letter, which accused Patton of negligence. “The way we heard it in class was indicative of a much more hurtful word with tremendous implications for the Black community.” The video of the class has since been circulated widely online, even spreading on Chinese social media, with Chinese viewers defending Patton’s use of the phrase and expressing confusion about why it was viewed as problematic. (Yeung, 2020)

⁶ See also Ethier (2020) about the verdict of the university after an internal investigation where Patton was found to have done no wrong.

Patton was replaced as the lecturer of that course and later apologised for causing “discomfort and pain”. He said that he had not realised “this negativity” or would have used a different example. However, some said that the idea that ‘*nèi ge*’ is offensive only makes sense from the perspective of English speakers, but that the Chinese language “is not subject to English rules”. If Chinese is not subject to English rules, why would Patton be removed from his class, and why would he feel the need to apologise?

A similar controversy, but at a much larger scale, occurred in 2019 when the K-Pop band BTS released their song “Fake Love” which included the Korean word for ‘mine’, read ‘*naega*’. The group eventually released another version for the US market, a decision which KJ103 Oklahoma City radio station host JJ Ryan explained:

My understanding is that they did this to avoid stations potentially not playing the single because it could be considered offensive OR generate complaints with the FCC. If the public finds anything on US radio “offensive” people can file a formal complaint. Those complaints COULD result in the stations broadcast license being revoked.⁷

As in the previous cases, many pointed out that there are more languages in the world besides English, and that “it’s disrespectful to censor a Korean word just because it sounds like an English slur.” (Kelly, 2018) A reply to JJ Ryan’s tweet by Nana (@GALAXYSTAR_98) said, “The public is going to assume [BTS] really are cussing [sic] and saying the N word now.”

To recapitulate, the cases can be categorized into four different types of misunderstanding.

- (1) Mentions or fictional utterances that are confused with derogatory uses.
- (2) Utterances with words with different etymologies, and different meanings, within the same language.
- (3) Utterances with words that share an etymology, but that have a different meaning, belonging to different languages.
- (4) Utterances of expressions with different etymologies, different meanings, and belonging to different languages.

The reaction from audiences to these different types of case is common, in that audiences mistaked the utterances for derogatory ones, were offended by what they thought the speaker did, and expected the speaker to be punished, to apologise, and to retract. Nonetheless, the responses from the speakers were more varied. In one case, the institution that employed the speaker issued a genuine apology for the (alleged) use of a slur (Mr. Collins case). In another, a speaker resigned and admitted to “having learned his lesson” (David Howard’s case). In a third case, the speaker accepted the punishment, apologised, and retracted (Cavani’s case). A different case had a speaker issuing an apology that was not genuine (Prof. Paton’s case). Finally, there was a clear retraction without any apology (BTS’s love song case).

I think that most cases involved a retraction, but not all were meant as retractions of *slurs*. In the last section, I argue that genuine apologies for slur use are indirect retractions of derogatory illocutionary acts. As a result, the cases here considered can

⁷ Ryan made this statement on Twitter: <https://twitter.com/JJRyanOnAir/status/997604005563060230?s=20&t=—IrhQE4XZcJJTJpem24TQ>.

be divided into two groups. Two were intended as retractions of slurs. Two others were not so intended. And one case was less clear. In the next section, I introduce some of the main current views about slurs before I finally answer, in the final section, the two remaining questions.

2 Slurs and expressives

The cases in the previous section involved a specific kind of error. Hearers mistook an utterance for a slurring derogation. It was because of this error that the audience was outraged and offended. Therefore, we should explain what it seemed to hearers that the speaker had said or done that explains the reasonableness of the hearers' conduct. This, in turn, requires an explanation of derogation.

I will not offer an exhaustive review of all the extant theories of slurs and pejorative terms, but merely of what I take to be the more salient or promising accounts. Others have provided independent criticism of some of these views. I will further aim to explain the rationality and reasonableness of feelings of anger, offence, or outrage in response to actual acts of derogation. This will help explain why audiences who *think* they heard slurs being used reacted with the same offence as if they had witnessed an actual derogation.

In this section I start with a brief and non-exhaustive overview of truth-conditional and propositional theories, and point to some of their limitations, followed by an introduction of use-conditional expressive views.

2.1 A non-exhaustive overview of theories of slurs, and some considerations

Some views about slurs identify their meaning as part of their truth-conditional semantic content (e.g., Hom, 2008, 2010, 2012; Hom & May, 2013, 2018). Others explain slurs' meaning in terms of use-conditions rather than truth-conditions. Use-conditional accounts are often cashed out in terms of expressive presuppositions (Macià, 2002, 2014; Schlenker, 2007; Cepollaro, 2015; Cepollaro & Stojanovic, 2016). Some use-conditional expressive views explain the content of expressive presuppositions with input from contemporary philosophy and psychology of emotions (Marques & García-Carpintero, 2020). Other accounts focus on conventional implicatures (Potts, 2007a, 2007b; Williamson, 2009), or on non-presuppositional expressive content (Jeshion, 2013a, 2013b, 2016; Richard, 2008). Finally, some authors argue for speech-act accounts of derogatory language (Langton, 2012; Bianchi, 2014; Camp, 2018). Some of the latter views are probably compatible with each other, although I won't argue the point here.

Other authors offer pragmatic accounts that don't look to the meaning of slurs, but rather to socio-pragmatic aspects of their use. Jorgensen Bolinger (2017) advocates for a contrastive account of slurs, where a speaker who chooses to use a slur rather than a neutral counterpart signals the speaker's endorsement of the term and of its associations (Jorgensen Bolinger, 2017). Similarly, Nunberg (2018) argues that "slurs function by evoking attitudes about the target that are associated with the group who have

constituted the word's historical provenance" (2018, p. 285). Anderson and Lepore (2013) claim that slurs are prohibited words, and it is because of this that using a slur produces offence. Diaz-Legaspe (2020) proposes that slurs are words with a specific kind of register.

Some of the accounts mentioned face substantial objections. Many have pointed out that the derogatory content of slurs appears to project out of semantic embeddings (under negation or conditionals, in different sentential moods, under possibility modals, etc.), a feature that truth-conditional theories struggle to explain (Jeshion, 2013a, pp. 316–319; Camp, 2018, p. 35; Cepollaro & Thommen, 2019; Sennet & Copp, 2015; Marques, 2021a). Moreover, Rappaport (2019) shows that semantic or content-based theories fail to meet what he calls the *Principle of Substitutivity Salva Iniuria* (PSSI).

[I]f two expressions or phrases have the same linguistic meaning, then in any context, they can be substituted for one another and produce the same degree of offense. (Rappaport, 2020, p. 178)

Truth-conditional theories thus appear unable to capture the *toxicity* of slurs, that is, the strong negative emotions of the speaker and the strong emotional reaction of outrage *elicited* or *evoked* from hearers. This has implications for the kind of audience error that concerns us here. The problem is not just that slurs project out of diverse sentential constructions and sentential modes. It is also that words that *are not* slurs appear to evoke the same degree of offense from audiences who disapprove of the slur's use. Accounts that struggle to explain the *iniuria* characteristic of slurs can be expected to also struggle to provide an explanation of the cases of miscommunication that concern us here.

Some authors see this as evidence that the offensiveness of slurs is not part of their semantic profile but is merely a socio-pragmatic effects. Nunberg (2018) emphasizes that an account of slurs must capture their normative force as words that serve to disparage people based on their group membership. Slurs, he contends, are just a special case of the way people explore sociolinguistic variation to display social status. By using a slur rather than a neutral counterpart referring to the same group, the speaker signals his group affiliation. (Jorgensen Bolinger (2017) has a similar but slightly different account, in terms of a speaker endorsing the associations of the word used.) These views point to sociolinguistic mechanisms that are independently present in communication. However, they are not without problems. First, not all slurs have neutral counterparts (Jeshion, 2013a, 2013b: 310). Indeed, there are languages where the slur for a group is also the noun used by group members to refer to themselves.⁸ Second, and more importantly, social accounts face the same difficulties that affect other views. Nunberg objects to theories of slurs which explain their meaning via the expression of a belief or attitude, typically *of the speaker*, towards the target of the slur. As he says:

⁸ See for instance Zeman (2020) on the Romanian 'țigan', and Castroviejo et al (2020) on the Spanish 'gitano'.

Why should I care about the attitudes of some pseudonymous bozo who rails about the kikes in a Twitter post—why should it matter to me or to anyone what he thinks or feels about Jews? (Nunberg, 2018, p. 284)

However, and as Marques and García-Carpintero (2020, p. 143, fn.13) point out, Nunberg’s own theory faces the same problem: Why should we care about the attitudes of a *group* of “pseudonymous bozos”, if we don’t care about what *one of them* thinks or feels? The same problem affects Jorgensen Bolinger’s proposal. Why should we care if a speaker endorses the associations of the term?

Anderson and Lepore (2013) offer a different explanation. For them, a slur’s offensiveness is to be construed entirely in terms of the associated social taboo. They consider one of the cases discussed in this paper, “niggardly”, and say: “the mere phonetic and orthographic overlap caused as much a stir as more standard offensive language” (2013, p. 23, note 29). They allege that content-based accounts are “useless in this regard”. But this is a *non sequitur*. Just as an externalist about epistemic warrant must explain how a subject’s conduct can be rationally explained despite the lack of *evidence*, by externalist lights, so too must semantic accounts explain people’s behaviour based on what they claim their interlocutor has said. Anderson and Lepore’s solution to Howard’s use of “niggardly” (and the remaining cases considered in this paper) is also disputable. First, “mere phonetic and orthographic overlap” surely explains *why* it seems to hearers that the speaker has used a slur, just as seeing someone with the same built and hair colour as my best friend in the distance explains why I call her name and wave, to greet her as a good friend should. But it’s not true that the use of the word “caused as much stir as *more standard* offensive language” would have caused. Had Howard actually used the n-word, we would not have seen the NAACP’s Chairman publicly criticise people’s “lack of understanding”, and the “stir” would have been considerably worst. Second, it’s questionable that the prohibitionist theory can explain the rationality of taking offence at derogation. The fact that a behaviour is prohibited need not make it *reasonable* or *fitting* to be offended by disobedience. Women in Iran are prohibited to remove their headscarves in public. That does not *warrant* offence against women who dare to disobey (see Jeshion, 2016, p. 138). In contrast, outrage at derogation *is right* and *warranted*. Taboo accounts cannot draw the line between the two. Third, and more importantly, *there was no pre-existing taboo against the use of “niggard”*, or against Chinese or Koreans or Uruguayans speaking their languages. A normative prescription regarding X does not warrant prescribing the same conduct towards things that look like X (as my greeting someone I mistake as my best friend illustrates). Therefore, Anderson and Lepore’s prohibitionism cannot account for the kind of cases discussed here. I turn now to use-conditional and expressive accounts.

2.2 Expressives and presuppositions

Is there hope that an expressivist account can do better? After Kaplan’s (1999) introduction of use-conditions, authors like Potts sought to identify common features of expressives. Despite claims that the offensiveness of slurs is not related to their semantic properties, Potts has convincingly argued that sentences with expressives

like ‘damn’ have conventional secondary meanings that are “semantically separate from the at-issue content but interact with it pragmatically” (Potts, 2012, p. 2516). In other words, (3) below doesn’t merely assert *that the dog is on the sofa*. It also communicates that the speaker “holds the dog in low regard” (Potts, 2007b: 171).

(3) The *damn* dog is on the sofa.

Potts (2007a) has also argued that expressives⁹ share six important features. Expressive content contributes to a dimension of meaning that is *independent* of its regular descriptive or truth-conditional content. That the dog is on the sofa is part of the at-issue-content, not that the speaker has a negative attitude towards the dog. Expressive content is *nondisplaceable*: when expressives occur embedded under various sentential constructions like negations or conditionals, with attitude verbs, modal operators, and in different sentential moods, their expressive content *projects*, a feature mentioned earlier. Expressives also have *descriptive ineffability*, i.e., expressive content cannot be fully paraphrased in descriptive non-expressive terms. This feature is manifest in the fact that truth-conditional propositional contents fail Rappaport’s PSSI. Finally, expressives have two additional features: *immediacy*—the intended expressive act is achieved immediately (as happens with a performative of the form “I promise you to...”), and *repeatability*—the repetition of the expressive reinforces the emotive content, rather than make it redundant or less intense. Slurs have other properties that distinguish them from expressives in general. Unlike mere expletives, pejoratives and slurs target *kinds* of people. Pejoratives do so based on very general features or character traits, but slurs seem to apply to people because of their membership in groups identified by relevant social traits, such as sexual orientation, sex, religion, ethnic or racial group.

Now, there are at least two points of contention about the nature of the not-at-issue content of slurs. First, there’s the question whether expressives express an attitude *of the speaker*. Second, there’s the question of whether expressives have cognitive or conative content.¹⁰ The idea that slurs express or signify an attitude *of the speaker* faces an objection that Ayer (1946) raised to metaethical subjectivism: in saying that war crimes are bad the speaker is neither asserting nor presupposing that *he* disapproves of war crimes (Ayer, 1946: 114). And for a hearer to accept that war crimes are bad is not for the hearer to accept that the speaker disapproves of war crimes. This problem affects contemporary accounts of expletives and slurs that hold that a “correct and accurate” use of a slur presupposes (or implicates) that the speaker believes that the members of the group targeted by the slur are contemptible (Potts, 2007a; Schlenker, 2007).¹¹

As to the second point of contention, Marques and García-Carpintero (2020) rely on arguments developed in earlier work by García-Carpintero (2004, 2015) against what he calls the “flattening strategy”, i.e., the attempt to reduce the illocutionary force of a speech act to a proposition describing that act. There are good reasons against the

⁹ Perhaps also evaluatives. See e.g., Stojanovic (2012), Cepollaro and Stojanovic (2016).

¹⁰ I’m grateful to a reviewer for raising this issue.

¹¹ Others have argued that slurs and pejoratives carry a presupposition of commonality (e.g., Macià 2002, 2014; see Marques 2022 for discussion), or that what is taken for granted is that the target is bad in some way (Cepollaro and Stojanovic 2016, Williamson 2009).

flattening strategy in general, and substantive reasons against it in the particular case of slurs. On the general reasons, the authors say,

An assertion that a command is given or a question posed can occur without the command being given or the question posed. Conversely, the non-cognitive attitude/act (the command or the question) can occur without the cognitive one (the belief/assertion that the command or the derogation was made), for instance because the thinker/speaker lacks the conceptual resources to describe the noncognitive state/act. (Marques & García-Carpintero, 2020, p. 142).

Slurs are words that conventionally serve to perform derogatory illocutionary acts, they argue. The authors define derogation by means of a constitutive norm:

derogation: one must derogate group G on account of their having features $F1 \dots Fn$ only if group G has $F1 \dots Fn$, and contempt fits a group with such features. (Marques & García-Carpintero, 2020, p. 147)

This norm is not part of the content of what is asserted, accepted, or believed with a use of a slur. Nor is it part of a presupposed proposition. Rather, a “correct and sincere” use of a slur *presupposes contempt* for the group that is the target of the slur (i.e., the group with the presumed negative features). To accept a slurring utterance is, then, in part to accept the fittingness of contempt towards its target.¹²

Rappaport (2020) argues that any full account of slurs must explain what he calls *toxicity*: the fact that slurs are “used to express intense emotions, and to cause strong emotional reactions in hearers.” (p. 178). Rappaport’s *Principle of Substitutivity Salva Iniuria* (PSSI) is consistent with Marques and García-Carpintero’s argument—that one *can* derogate without having the conceptual resources to describe the noncognitive state/act. This is supported by the research in neurolinguistics cited by Rappaport. One can derogate without having the relevant conceptual resources, i.e., without the cognitive ability and semantic capacity to process complex propositional contents:

For the present purpose, however, the important point is that these findings are consistent with the claim that the forms of swear words have direct access to emotional processing centers in the brain, unmediated by higher-level cognitive analysis. This conclusion is supported by various behavioural studies that also suggest that stimuli can evoke affective responses independently of cognitive (semantic) analysis. (Bowers & Pleydell-Pearce, 2011, cited in Rappaport, *op.cit.*)

Since there is some evidence that the affective contents expressed with slurs are not processed in the same way as other cognitive, semantic, contents (Rappaport, 2019, 2020), this can help explain the descriptive ineffability of derogatory meaning: that

¹² Perhaps someone could advance a similar related account where a use of a slur presupposes that the speaker is issuing a directive issued to the audience: *to treat group G with contempt*. This would update a different part of Stalnakerian contexts to those updated by assertions. It’s not entirely clear if this solution would satisfy PSSI as formulated by Rappaport. After all, one can react to a use of a slur while lacking the cognitive resources needed to process that one is being directed to treat others in a derogatory manner. I thank a reviewer for raising this issue.

expressive content cannot be fully paraphrased in descriptive non-expressive terms (Potts, 2012). It also explain slurs' toxicity.

If the emotions or affects expressed are what account for the toxicity of slurs, then we must understand their role, particularly emotions like contempt.¹³ Jeshion (2013a) urges theorists to do so:

The nature of contempt infects slurring terms' semantics. Contempt involves taking those properties that are the basis for regarding the target contemptuously as fundamental to the targets' identity as a person, and this feature of contempt is encoded. (Jeshion, 2013a: 333)

Marques and García-Carpintero (2020) also claim that slurs presuppose *contempt*, drawing from work in moral psychology and the philosophy of emotions to understand the normative appraisals that characterize it: Emotions involve an *appraisal* of a stimulus (a situation, event, someone, or someone's conduct towards us), they *feel* a certain way (have a particular phenomenology), involve a particular type of *motivation* or *desire*, and tend to lead to certain *action tendencies*.

In the next and final section, I'll show how this model can help us answer the two questions formulated at the beginning of this paper.

3 Apologising and retracting for what one didn't do, didn't say, and didn't mean

The questions this paper set out to answer were:

- (i) Are all cases considered genuine instances of retraction?
- (ii) If the cases considered are not cases of slurs, in what sense (if any) is the paper still about the retraction of slurs?
- (iii) Do the cases support expressive accounts of slurs?

As I said at the outset, the second question has a rather straightforward answer, since it is standard in philosophy to reflect on perceptual errors, for instance, to make sense of the explanatory challenges such examples raise for theories of perception. What concerns us here is the conduct of people who mistake some non-derogatory utterances for acts of derogation. In what remains of this paper, I will first address question (iii), and then question (i), identifying which of the cases are genuine (infelicitous) retractions of derogatory acts.

3.1 Diagnosing the error, making sense of the offense

In Sect. 2, I presented reasons for an account where uses of a slur presuppose contempt towards the slur's target. I also pointed towards arguments others have made against truth-conditional semantic accounts and provided independent reasons against socio-pragmatic views' ability to deal with the cases of misunderstanding that concern us here.

¹³ See Mulligan (1998), D'Arms and Jacobson (2000), Mason (2003).

Let us assume that to use a slur is to (defeasibly) perform a strongly derogatory act, which presupposes contempt towards the target group. Now, being the target of contempt can elicit anger (or variations of anger, like indignation or outrage). Indeed, anger, contempt, and disgust are reactive attitudes (cf. Strawson, 1962) and are characterized by different appraisal norms:

These three emotions are elicited by *violations of different moral codes*: (1) anger is triggered by violations of individual rights and autonomy and appraisals of self-relevance; (2) contempt by violations of communal codes and hierarchy and appraisals of other's incompetence or lack of intelligence; and (3) disgust by violations of codes for purity and sanctity and appraisals of other's moral untrustworthiness. (Hutcherson & Gross, 2011; Rozin et al., 1999; Matsumoto et al., 2015, p. 2. My emphasis)

If slurs are conventional means of performing derogatory acts via the presupposition of contempt against members of a group, and contempt appraises its target as (socially or morally) inferior *qua* people, then *anger* at expressions of contempt is itself *apt*. This would be the explanation of why it is rational to take offence at derogatory utterances.

Reactive emotions are appraisal states elicited by norm violations, and they are themselves normative states, in a sense. Thus, slur use, if accepted, enacts the appraisal norms typical of contempt, i.e., appraises the members of the group that is derogated as deserving contempt for being inferior, immoral, incompetent, etc. In turn, being offended by derogation is a way of being outraged or angry at the injurious nature of contempt, since anger is apt in response to acts which are *injurious*. In other words, A's contempt for B appraises B as inferior, unworthy, or untrustworthy, and B's anger at A's contempt appraises it as a violation of B's self-worth.

What about the cases of utterances mistaken for derogatory? How to explain the reaction of the audience that is angered, but mistaken? Their reaction would have been reasonable, and in fact fitting, had the speaker used a slur. But a mistake in the interpretation of what a speaker said or did can evoke the same indignation as a correct interpretation, even if it is unfitting in the cases at hand. Anger or outrage is not fitting in the present cases because the speaker did not derogate the audience's "appraisal of self-relevance". Therefore, the cases considered here are consistent with expressive accounts of the meaning of slurs that properly take into account the moral psychology of the emotions expressed. The difference between the retraction of constatives and of expressives further supports this result.

3.2 Kinds of apology, punishment, and retraction

The only question left to answer is whether all the cases in the first section should count as retractions of slurs (or of derogatory acts), exploring the clues the examples provide concerning the retraction of slurs as the undoing of the normative appraisal of contemptuous derogation. The examples of Sect. 1 fall into two different groups, as I now argue. First, some are clear examples of intended retractions of slurs. Second,

others are cases that were not performed as retractions of slurs. Finally, there is one case that is less clear, but I think it should count as a genuine intended retraction.

As I said at the outset, I agree with Caponetto's (2020) Austinian account of retraction. A retraction of a speech act previously performed undoes the normative commitments enacted by that previous act. It is to be expected, given this, that a retraction of a *constative* act such as an assertion undoes commitments that are distinct from those enacted by an *expressive* act. As Caponetto points out, retractions only undo *the normative commitments* of past speech acts, but do not erase the past locution from the conversational record. As speech acts themselves, retractions have felicity conditions, which Caponetto lays down as follows:

General felicity conditions of retraction: a speech act A performed by a speaker S at a time t may be retracted at a later time t_n only if

- i. A was felicitously performed at t ;
- ii. (a) A was performed by the retractor, or
 - (b) the retractor is a third party whose position grants her the right to take back A . (Caponetto, 2020, pp. 2409–2410)

Several things are of note here. If the original act was *not* felicitously performed, then its retraction will not be felicitous either. Second, some people other than the speaker may be in a position to retract the original speech act. Presumably, this occurs when academic journals retract papers even though such papers were not authored by the editorial board of the journal.

Now, authors like MacFarlane (2014) have indicated that when one is shown that an assertion one made is not true (2014: 108), one would say things like “I guess I was wrong” (2014: 240) or “Scratch that. I was wrong.” Since standard presuppositions also add propositions to the common ground, we can assume that presuppositions can be retracted in similar ways with “I guess I was wrong”. Let us now continue to assume that slurs expressively presuppose contempt for their target and (defeasibly) perform derogatory acts. If this is so, it is expectable that the response to a presupposition of contempt differs from a response to a presupposition of a false proposition. Contrast the two examples (4) and (5) below:

- (4) A. It was John who finished all the biscuits. [presupposition: someone finished the biscuits].
 - B. What are you talking about? There are plenty of biscuits left in the tin!
 - A. Oh, I guess I was wrong. I take that back.
- (5) A. The *charnegos* next door are celebrating Real Madrid's win [presupposition: contempt for Spanish speaking Spaniards in Catalonia].
 - B. That is very rude! You can't speak of people like that!
 - A. I'm sorry, I won't do it again.

These are plausible reconstructions of possible interactions. If slurring utterances perform derogatory speech acts, their retraction should undo the deontic changes introduced by derogation, in this case, the presupposition of (fitting) contempt against the target of the slur. Merely saying “I take that back” is not enough. On the other hand, if one asserts something that is not true, one is required to retract, but not to

issue a *genuine apology*. A genuine apology is owed only when one does something morally wrong (Bovens, 2008), and one is further required to atone or make amends. This puts pressure on attempts to reduce the content of slurs to truth-conditional, propositional, contents. Merely saying “I take that back” cannot make amends, and does not acknowledge the wronging of derogation, any more than returning stolen money suffices to make amendments for theft.

But consider now how A would respond if B misunderstood some of her presuppositions:

(6) A. It was John who finished all the biscuits.

B. What are you talking about? There are plenty of biscuits left in the tin!

A. Yeah, but I didn't mean *those* biscuits. I meant the new packet that I bought this morning!

B. Oh, I thought you meant the other ones.

A's utterance carried an additional presupposition concerning the domain restriction of 'the biscuits.' In such cases, speakers in B's position can (and should!) just accept that they misunderstood the speaker A's intentions. A neither retracts nor apologises, since what A said, and presupposed, is true and warranted. This strongly contrasts with the behaviour of audiences in the cases presented in the first section. The controversy does not end when speakers clarify what they said, meant or intended to say. Not only that, as I will now try to show, at least *some* of the retractions discussed earlier should count as intended retractions of derogatory acts, even if infelicitous ones.

To recall, here's how speakers acted in the face of their audience's challenge:

- The school apologised for Mr. Collins behaviour (“we apologised to students and families”).
- Edinson Cavani retracted (deleted the Instagram post), apologised, and accept the punishment from the F.A.
- Howard resigned (accepting punishment) and said that he had “learned [the lesson] from the situation.”
- Prof. Patton apologised “for causing” pain or offence.
- BTS retracted the song and released a different one (no apology).

If derogatory acts are expressive of contempt, they are morally problematic. Hence, a genuine apology that admits moral blame is owed, in order to atone for the wrongful action. My conjecture is that apologies for wrongful speech indirectly perform retractions, as the toy example in (5) (“I'm sorry. I won't do it again”) suggests. But there are different kinds of apologies.

Some apologies are expressions of regret *that something happened*. “I'm sorry your wallet was stolen” expresses regret that something bad happened to you, and perhaps my sympathy, but I'm not to blame. Maybe I could have contributed causally to it, say, because I entertained you until late at night and you ended up taking public transport at an hour that made it more likely that you'd be robbed. Still, I'm not morally responsible for the theft. And so *I don't owe you a genuine apology*.

Now, some of the speakers in the cases described earlier apologised. *Some* of the apologies issued in these cases seem to accept moral responsibility, not merely causal responsibility, or regret *that* something bad happened. Apologising and retracting is a

way of accepting moral responsibility and the “punishment” that is due: A’s contempt for B (expressed through a derogatory act) appraises B as inferior or unworthy, and B’s anger at A’s contempt appraises that contempt as a violation of B’s self-worth. As Matsumoto et al. (2015) put it, “anger is triggered by violations of individual rights and autonomy and appraisals of self-relevance.” Callard’s (2020) insightful analysis of anger further illuminates the nature and logic of retribution:

Revenge allows you to turn the principle of my action into a rule for your conduct toward me: you make my bad your good. This is the opposite of trying to undo or reverse my action. You hold me accountable by holding onto my theft, refusing to forget it, turning its one-off opposition between our interests into a rule to which I am now subject. You do not let me “live it down,” instead you force my own thinking down my throat. Seeing *me as accountable for what I have done* means treating my action as a principle governing our interactions. Angry people sometimes describe their vengeance as “*teaching someone a lesson*,” and this is quite literally true: you make my wronging of you into a general principle and then “educate” me by imposing it on me. (Callard, 2020. My emphasis).

A better understanding of the moral psychology of anger and of the function of genuine apologies thus provides the rationale for relating genuine apologies and the retraction of derogatory acts. One cannot “undo the action”, in Callard’s words. That is, one cannot undo the perception of the insult, and the outrage it elicits. This is so whether an actual derogation/offence took place or not—one cannot undo the perlocutionary effects of an utterance. Moreover, speakers responsible for derogatory speech will be held accountable for it and will be required to retract; but to retract is not to undo perlocutionary effects. It is to cancel the deontic ones.¹⁴ The undoing of the deontic wrong of derogation plausibly requires a genuine apology, as a sign of accepting moral blame, i.e., as an indication that one has “learned a lesson”. Hence, behaviour that manifests atonement in this fashion is plausibly indicative of “undoing the commitments enacted by the previous act”, that is, of its retraction.

This allows us to count David Howard’s case (who used “niggardly”) as the intended retraction of a derogatory act, even if he performed no such thing. He accepted the blame first by resigning his position, and then by saying that he “learned from the situation”. Two of the other cases are more straightforward: Cavani’s (he deleted his Instagram post and accepted the punishment that the FA imposed), and Mr. Kingsley, headteacher at Mr. Collins’s school, who apologised to parents and students (on behalf of the school). Although it was not Mr. Collins who apologised, the case still meets one of Caponetto’s conditions of felicitous retractions: “ii. b. the retractor is a third party whose position grants her the right to take back A.” Accepting blame, apologising, admitting to “having learned a lesson”, are part of what a retraction of wrongful speech would require. This is all compatible with performing such intended retractions *infelicitously*, because they are not retractions of actual derogations, since none were performed in the first place, arguably.

A different set of cases that do not appear to have been intended as retractions of derogatory acts are that of Prof. Patton, who apologised *for* causing offense, or

¹⁴ Chris Bennet (2008) develops a theory of punishment based on the practice of (genuine) apology.

BTS's case, that simply retracted the song and launched a similar one with changed lyrics. Patton's case illustrates the difference between a genuine apology for a moral wrong, and a mere expression of regret *that* something happened (which may be more pertinent when one plays a merely causal role). He did not retract, and indeed was found to have done nothing wrong by his university. The K-Pop band *retracted* so as not to cause misunderstandings, and then released a different song. This case should count as a retraction as much as a retraction of a paper from a journal does.

4 Conclusion

This paper considered cases of morally loaded *seemings*: non-derogatory utterances that seem to certain audiences to be otherwise. The cases raised three questions. are the cases considered genuine retractions? If the speakers didn't derogate by using a slur, how are the cases retractions of *derogatory acts*? Do these examples support expressive accounts of slurs?

To answer these questions, I had to appraise contemporary accounts of slurs and derogatory language. I briefly reviewed semantic theories that explain the meaning of slurs in terms of *propositional* contents and summarized some of the standing objections against them. I also reviewed some socio-pragmatic accounts that could aspire to offer better explanations of the use of slurs. I questioned their capacity to explain regular slur use and argued that they cannot accommodate the cases considered in this paper. I further provided reasons to see slurs as a special kind of expressive with a distinct normative impact. Accepted slur use normalizes regarding the target of the slur with contempt. It is the assumption that this contempt is fitting that warrants audiences' reaction of moral outrage. It is thus the fact that *it seems* to the audience that a slur was used that explains the audience's reaction. Finally, I argued that retracting a derogation requires more than *saying* that one retracts; it also requires an apology for the wrong of derogation. In the cases here considered, genuine apologies, which acknowledge moral blame, are not owed. However, people often do what they need not do. The retractions that were intended as retractions of derogatory acts are infelicitous. Nonetheless, the fact that at least some speakers made genuine apologies and retractions, and that audiences demanded them, reveals a specific kind of linguistic injustice.

Acknowledgements Earlier versions of this paper were presented at the Retraction Workshop organized by the University of Warsaw, Poland, in June 2021, and at the 23rd Amsterdam Colloquium in December 2022. I'm grateful to the audiences at those events. Special thanks to Dan Zeman, Grzegorz Gaszczyk, Craig Roberts, Frank Veltmann, Manuel García-Carpintero, and anonymous reviewers for this journal for their useful comments and criticism. Work towards this paper was financially supported by grant CEX2021-001169-M funded by MCIN/AEI/<https://doi.org/10.13039/501100011033>, by the project PID2020-119588GB-I00 of the Spanish Ministry of Science and Innovation, and by 2021-SGR-00276 for the Consolidated Research Group LOGOS, by the Generalitat de Catalunya.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. This work is supported by the Ministerio de Ciencia, Innovación y Universidades, PID2020-119588 GB-I00, Teresa Marques, Ministerio de Ciencia e Innovación, CEX2021-001169-M, and by 2021-SGR-00276 for the Consolidated Research Group LOGOS, by the Generalitat de Catalunya.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Academia Nacional de Letras de Uruguay (2020, Dec 31). Declaración de la Academia Nacional de Letras del Uruguay ante la sanción a Edinson Cavani. Retrieved from <http://www.academiadeletras.gub.uy/innovaportal/v/124868/46/mecweb/ante-la-sancion-a-edinson-cavani>.
- Anderson, L., & Lepore, E. (2013). Slurring words. *Noûs*, 47(1), 25–48.
- Ayer, A. J. (1946). *Language, truth and logic* (revised). Victor Gollancz.
- Bennett, C. (2008). *The apology ritual: a philosophical theory of punishment*. Cambridge University Press.
- Bianchi, C. (2014). The speech acts account of derogatory epithets: Some critical notes. In J. Dutant, D. Fassio, & A. Meylan (Eds.), *Liber Amicorum Pascal Engel* (pp. 465–480). Université de Genève.
- Bovens, L. (2008). Apologies. *Proceedings of the Aristotelian Society*, 108, 219–239. <https://doi.org/10.1111/j.1467-9264.2008.00244.x>
- Bowers, J. S., & Pleydell-Pearce, C. W. (2011). Swearing euphemisms, and linguistic relativity. *PLoS ONE*, 6(7), e22341. <https://doi.org/10.1371/journal.pone.0022341>
- Callard, A. (2020). On anger. *Boston Rev.* Forum 13 (45.1).
- Camp, E. (2013). Slurring perspectives. *Anal Phil*, 54(3), 330–349.
- Camp, E. (2018). A dual act analysis of slurs. In D. Sosa (Ed.), *Bad words* (pp. 29–59). Oxford University Press.
- Caponetto, L. (2020). How to undo things with words. *Synthese*, 197, 2399–2414.
- Castroviejo, E., Fraser, K., & Vicente, A. (2020). More on pejorative language: insults that go beyond their extension. *Synthese*, 198(10), 9139–9164.
- Cepollaro, B. (2015). In defense of a presuppositional account of slurs. *Language Sciences*, 52, 36–45.
- Cepollaro, B., & Stojanovic, I. (2016). Hybrid evaluatives: in defense of a presuppositional account. *Grazer Philosophischen Studien*, 93, 458–488.
- Cepollaro, B., & Thommen, T. (2019). What's wrong with truth-conditional accounts of slurs. *Linguistics and Philosophy*, 42(4), 333–347. <https://doi.org/10.1007/s10988-018-9249-8>
- D'Arms, J., & Jacobson, D. (2000). The moralistic fallacy: on the “appropriateness” of emotions. *Phil Phenomenol Res*, 61(1), 65–90.
- Diaz-Legaspe, J. (2020). What is a slur? *Philosophical Studies*, 177(5), 1399–1422. <https://doi.org/10.1007/s11098-019-01259-3>
- Ethier, M. (2020, 26, Nov). USC Marshall finds students were sincere, but prof did no wrong in racial flap. *Poets and Quants*. Retrieved from <https://poetsandquants.com/2020/09/26/usc-marshall-finds-students-were-sincere-but-prof-did-no-wrong-in-racial-flap/>
- García-Carpintero, M. (2004). Assertion and the semantics of force-markers. In C. Bianchi (Ed.), *The semantics/pragmatics distinction* (pp. 133–166). CSLI Publications.
- García-Carpintero, M. (2015). Contexts as shared commitments. *Frontiers in Psychology*, 6, 1–13.
- García-Carpintero, M. (2019). Conventions and constitutive norms. *Journal of Social Ontology*, 5(1), 35–52.
- García-Carpintero, M. (2020). On the nature of presupposition: A normative speech act account. *Erkenntnis*, 85(2), 269–293.
- Hom, C. (2008). The Semantics of racial epithets. *Journal of Philosophy*, 105(8), 416–440.
- Hom, C. (2010). Pejoratives. *Philosophy Compass*, 5(2), 164–185.
- Hom, C. (2012). A puzzle about pejoratives. *Philosophical Studies*, 159(3), 383–405.
- Hom, C., & May, R. (2013). Moral and semantic innocence. *Analytical Philosophy*, 54(3), 293–313.
- Hom, C., & May, R. (2018). Pejoratives as fiction. In D. Sosa (Ed.), *Bad Words*. Oxford University Press.

- Hutcherson, C. A., & Gross, J. J. (2011). The moral emotions: A social–functionalist account of anger, disgust, and contempt. *Journal of Personality and Social Psychology*, 100(4), 719–737. <https://doi.org/10.1037/a0022408>
- Jackson, J. (2020, Nov 11). Edinson Cavani apologises for using racial term in Instagram post. *The Guardian*. Retrieved from <https://www.theguardian.com/football/2020/nov/30/manchester-uniteds-edinson-cavani-apologies-for-racist-instagram-post>.
- Jeshion, R. (2013a). Expressivism and the offensiveness of slurs. *Philosophical Perspectives*, 27(1), 231–259.
- Jeshion, R. (2013b). Slurs and stereotypes. *Analytical Philosophy*, 54(3), 314–329.
- Jeshion, R. (2016). Slur creation, bigotry formation: The power of expressivism. *Phenomenology and Mind*, 11, 130–139.
- Jorgensen Bolinger, R. (2017). The pragmatics of slurs. *Noûs*, 51(3), 439–462.
- Kajumba, A. (2020, Nov 11). ‘The last thing I wanted to do was cause offence to anyone’: Edinson Cavani apologises for calling a friend ‘negrito’ on Instagram and says he’s ‘completely opposed to racism’ as FA launch probe which could see him hit with MINIMUM three-game ban. *Daily Mail*. Retrieved from <https://www.dailymail.co.uk/sport/sportsnews/article-9102391/Manchester-striker-Edinson-Cavani-handed-three-match-ban-FA.html>.
- Kaplan, D. (1999). *The meaning of ‘ouch’ and ‘oops.’* MS.
- Kelly, E. (2018, May 22). Here’s why BTS’s Fake Love is being censored on US radio. *Vanity Fair*. Retrieved from <https://metro.co.uk/2018/05/22/btss-fake-love-censored-us-radio-7568200/>.
- Langton, R. (2012). Beyond belief: Pragmatics in hate speech and pornography. In M. K. McGowan & I. Maitra (Eds.), *Speech and harm: Controversies over free speech* (pp. 72).
- Macià, J. (2002). Presuposición y significado expresivo. *Theoria*, 17(3), 499–513.
- Macià, J. (2014). Expressive meaning and presupposition. *Handout for a talk at the names, demonstratives, and expressive conference in Gargnano, Italy*.
- Marques, T. (2015). Retractions. *Synthese*, 195, 3335–3359.
- Marques, T. (2016). Pejorative discourse is not fictional. *Thought: A Journal of Philosophy*, 4, 1–14.
- Marques, T. (2021a). Pejoratives & oughts. *Philosophia*, 49(3), 1109–1125.
- Marques, T. (2021b). Illocutionary force and attitude mode in normative disputes. *Metaphilosophy*, 52(3–4), 1–17.
- Marques, T. (2022). The expression of hate in hate speech. *Journal of Applied Philosophy*, 10, 1–29.
- Marques, T., & García-Carpintero, M. (2020). Really expressive presuppositions and how to block them. *Grazer Philosophischen Studien*, 97(1), 138–158.
- Mason, M. (2003). Contempt as a moral attitude. *Ethics*, 113(2), 234–272.
- Matsumoto, D., Hwang, H. C., & Frank, M. G. (2015). The role of intergroup emotions in political violence. *Current Directions in Psychological Science*, 24(5), 369–373.
- MacFarlane, J. (2014). *Assessment sensitivity: Relative truth and its applications*. Oxford: Oxford University Press.
- Mulligan, K. (1998). From appropriate emotions to values. *The Monist*, 81(1), 161–188.
- Nunberg, G. (2018). The social life of slurs. In D. Fogal, D. Harris, & M. Moss (Eds.), *New work on speech acts*. Oxford University Press.
- Potts, C. (2007a). The expressive dimension. *Theoretical Linguistics*, 33(2), 165–198.
- Potts, C. (2007b). Into the conventional-implicature dimension. *Philosophy Compass*, 2(4), 665–679.
- Potts, C. (2012). Conventional implicature and expressive content. In C. Maienborn, K. von Heusinger, & P. Portner (Eds.), *Semantics: An international handbook of natural language meaning*. De Gruyter Mouton.
- Rappaport, J. (2019). Communicating with slurs. *Philosophical Quarterly*, 69(277), 795–816.
- Rappaport, J. (2020). Slurs and toxicity. *Grazer Philosophische Studien*, 97(1), 177–202.
- Richard, M. (2008). *When truth gives out*. Oxford: Oxford University Press.
- Robinson, J. (2016, Aug 3). What *The People v. O.J. Simpson* got very wrong about that Mark Fuhrman N-word scene. *Vanity Fair*. Retrieved from <https://www.vanityfair.com/hollywood/2016/03/mark-fuhrman-n-word-people-v-oj-simpson-f-lee-bailey>.
- Rozin, P., Lowery, L., Imada, S., & Haidt, J. (1999). The CAD triad hypothesis: A mapping between three moral emotions (contempt, anger, disgust) and three moral codes (community, autonomy, divinity). *Journal of Personality and Social Psychology*, 76(4), 574–586. <https://doi.org/10.1037//0022-3514.76.4.574>
- Schlenker, P. (2003). A plea for monsters. *Linguistics and Philosophy*, 26(1), 29–120.

- Schlenker, P. (2007). Expressive presuppositions. *Theoretical Linguistics*, 33(2), 237–245.
- Sennet, A., & Copp, D. (2015). What kind of a mistake is it to use a slur? *Philosophical Studies*, 172(4), 1079–1104.
- Stojanovic, I. (2012). Emotional disagreement: The role of semantic content in the expression of, and disagreement over, emotional values. *Dialogue*, 51(1), 99–117.
- Strawson, P. (1962). Freedom and resentment. *Proceedings of the British Academy*, 48, 187–211.
- Wheeler, C. (2020, Dec 12). Man United striker Edinson Cavani is handed a three-match ban by the FA for 'gracias negrito' Instagram post... meaning Uruguayan will return in time for crunch Liverpool showdown. *Daily Mail*. Retrieved from <https://www.dailymail.co.uk/sport/sportsnews/article-9102391/Man-United-striker-Edinson-Cavani-handed-three-match-ban-FA.html>.
- Williamson, T. (2009). Reference, inference, and the semantics of pejoratives. In *The Philosophy of David Kaplan* (pp. 137–158). Oxford University Press.
- Williamson, T. (1996). Knowing and asserting. *Philosophical Review*, 105(4), 489.
- Williamson, T. (2019). Morally loaded cases in philosophy. *Proceedings and Addresses of the American Philosophical Association*, 93, 159–172.
- Woodlee, Y. (1999, Feb 4). D.C. Mayor acted 'hastily,' will rehire aide. *The Washington Post*. Retrieved from <https://www.washingtonpost.com/wp-srv/local/longterm/williams/williams020499.htm>.
- Woolcock, N. (2021, June 21). Teacher used n-word in lesson on how not to pronounce Niger. *The Times*. Retrieved from <https://www.thetimes.co.uk/article/teacher-used-n-word-in-lesson-on-how-not-to-pronounce-niger-h0cbph8p0>
- Yeung, J. (2020, Sept 10). USC professor under fire after using Chinese expression students allege sounds like English slur. *CNN*. Retrieved from <https://edition.cnn.com/2020/09/10/us/usc-chinese-professor-racism-intl-hnk-scli/index>
- Zeman, D. (2022). A rich-lexicon theory of slurs and their uses. *Inquiry: an Interdisciplinary Journal of Philosophy*, 65(7), 942–966.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.