# Slurs and Expressive Commitments

**Leopold Hess¹** 🆔

## Abstract

Most accounts of the derogatory meaning of slurs are semantic. Recently, Nunberg (2018) proposed a purely pragmatic account offering a compelling picture of the relation between derogatory content and social context. Nunberg posits that the semantic content of slurs is identical to that of neutral counterparts, and that derogation is a result of the association of slur use with linguistic conventions of bigoted speakers. The mechanism responsible for it is a special kind of conversational implicature. However, this paper argues that Nunberg's proposal suffers from technical and conceptual problems regarding the application of the concepts of implicature and convention. It is proposed that the implicature mechanism should be replaced by expressive commitment attribution (developed for expressives in Hess *Journal of Pragmatics, 129*,13–33 2018, 2019). The account advocated here preserves crucial insights of Nunberg's analysis (and incorporates elements of other pragmatic accounts, cf. Anderson and Lepore *Noûs 47*(1), 25–48, 2013a, *Analytic Philosophy 54*(3), 350–363, b; Bolinger *Nous, 51*(3), 439–462, 2017), giving them a firmer conceptual basis. It supports a distinction between derogation and offensiveness of slurs that improves on the one proposed by Hom and May (*Analytic Philosophy 54*(3), 293–313 2013), and sheds light on puzzling borderline cases: the use of slurs by ignorant speakers and controversies surrounding the status of some words as potentially slurs.

## 1 Introduction

Language can be used to communicate information, coordinate actions, express feelings and beliefs, and create art. It can also be used to harm, insult, diminish and humiliate other human beings. *Slurs* are derogatory epithets that are specialized tools for performing these harmful and denigrating functions. They refer to individuals *qua* members of a group defined by some socially salient characteristic such as race,

---

✉ Leopold Hess
leopoldhess@gmail.com

1    University of Warsaw, Warsaw, Poland

ethnicity, sexual orientation and religion, and they do so in a pejorative way, expressing negative attitudes, insinuating harmful stereotypes, or invoking discriminatory ideologies. Most slurs are explosively offensive, especially that they usually target underprivileged or minority groups. And yet, unlike swear words and expletives, they function grammatically as regular nouns and adjectives. Consider the following contrast:

(1)   Aiko is Japanese.
(2)   Aiko is a jap.

While (1) simply predicates group membership of Aiko, based on her nationality or ethnicity, (2) derogates her and by extension all Japanese people *qua* members of this group. One way to explain the contrast between (1) and (2) is to assume that the slur and its neutral counterpart differ in semantic content, broadly construed. This difference can be explained in a variety of ways: the pejorative meaning of a slur may be part of its truth-conditional content (Hom 2008; 2012; Hom and May 2013), a conventional implicature (Potts 2007; Williamson 2009; McCready 2010; Whiting 2013), a presupposition (Schlenker 2007; Cepollaro 2015; Cepollaro and Stojanovic 2016), expressive content (Richard 2008; Croom 2011; Hedger 2012; Jeshion 2013a), use-conditional content (Predelli 2013; Gutzmann 2015), or content contributed to two distinct speech acts (Camp 2018). The philosophical literature on slurs has largely focused on evaluating the advantages and disadvantages of these various accounts.

Recently, Geoffrey Nunberg (2018) has proposed a theory of slurs that breaks away with this common assumption and explicates the pejorative import of slurs in a non-semantic, thoroughly pragmatic way.[1] According to his theory, derogatory meanings of slurs emerge as a special kind of markedness implicatures, which he calls 'ventriloquisitic' implicatures, because in using a slur a speaker is as if imitating a way of speaking characteristic of bigots and racists. In this paper, I argue that while compelling in many respects, Nunberg's theory has significant weaknesses and the notion of ventriloquistic implicatures is ill-suited to account for some important properties of slurs. Accordingly, I propose to replace it with the notion of expressive commitment (cf. Hess 2018, 2019) which not only makes it possible to base Nunberg's main claims on a more secure footing but also illuminates some less discussed aspects of slur use.

In Section 2, I present the most important elements, and virtues, of Nunberg's implicature-based theory of slurs. In Section 3, I discuss two major problems that it faces. I propose to solve these problems by replacing the implicature mechanism with *expressive commitments*. This concept is introduced in Section 4, and applied to the case of slurs in Section 5. Finally, in Section 6, I discuss several further aspects of slur use that my proposed theory helps to illuminate.

---

[1] Strictly pragmatic theories of slurs have also been proposed by Anderson and Lepore (2013b) and Anderson and Lepore (2013a) and Bolinger (2017). I will mention both in Section 6; they can be considered as complementary with respect to Nunberg's and my own account. I focus on Nunberg because his theory covers some aspects of slur use that are otherwise ignored, especially their role in signalling the speaker's affiliation with discriminatory practices. See below, Section 2.

## 2 Nunberg's Proposal

The central idea of Nunberg's long and insightful article, aptly titled 'The Social Life of Slurs', is that a full understanding of the nature of slurs must begin with the observation that their linguistic properties are inextricably linked with socio-political factors and issues that determine who uses slurs and to what purpose.

Nunberg argues at some length against what he considers a reductive 'semanticism' in most existing approaches to slurs. Discussing his arguments in detail would require a separate essay, but I will mention two claims about slurs which Nunberg makes that are not easily accommodated in content-based theories, and therefore add to the explanatory scope of his theory. First, while semantic accounts focus on the nature and mechanisms of the offense and derogation inflicted by slurs on their targets, they are ill-suited to explain another very important function of slurs, which is to support the self-identification of slur users as both distinct from the target group and affiliated with a group of their own. This is another way in which slurs differ from other pejoratives and expressives: unlike calling someone a 'jerk' or 'dumbass', which serves mostly to express personal antipathy towards them, using a racist or homophobic epithet, Nunberg argues, functions not only to insult the target but also to signal the speaker's affiliation with a part of society they consider superior (e.g. whites as opposed to Blacks, citizens as opposed to immigrants, or straight people as opposed to homosexuals). This signal is addressed at the slur user's peers (and bystanders), and not just the targets of their abuse. It is an important aspect of what slurs are used for, and it explains why they are often used in the absence of any of the derogated people, in exchanges between racists, homophobes or anti-Semites.

The second claim concerns the status of the category of slurs. Nunberg argues that 'slur' is itself a thick term, i.e. a hybrid word that mixes categorization and attitude or evaluation, unlike, say, 'pejorative' or 'derogative'. We may consider a word derogative, but deem its use acceptable (sometimes assholes should be called 'assholes' and neo-Nazis, 'neo-Nazis'). But the use of a *slur* is morally objectionable. This cast doubts on the prospects of a semantic theory of slurs in general—if any attempt at delineating the class of slurs must involve *normative* considerations, it is prima facie implausible that they can be defined in purely semantic terms. Moreover, if 'slur' is in itself a thick, evaluative term, it is only natural that people may differ on whether a given word, such as 'Oriental' or 'fascist' is a slur (and hence unacceptable) or a descriptive (even if pejorative, like the latter example) term—and that those differences may often track differences of ideology or political positions. I will come back to this point in Section 6; the expressive commitment-based account I propose offers further elucidation of it.

Nunberg's own positive account places central focus on the association between expressions—slurs—and the kind of people—racists, homophobes etc.—who use them. As he summarizes it: 'In a nutshell: racists don't use slurs because they're derogative; slurs are derogative because they're the words that racists use'. The key observation here is that a slur is always just one possible way of referring to a given group; there are neutral, non-offensive alternatives. Unlike semantic theorists, Nunberg assumes that the encoded content of a slur is the same as that of its neutral counterpart. Against this background of alternatives, the derogatory meaning of slurs

arises as a *conversational implicature* through violation of the Maxim of Manner. These implicatures, which are routinized, are of what Nunberg calls a 'ventriloquistic' sort: the speaker employs a convention of a group other than the one he belongs to, thereby associating himself with this group and its (stereo)typical attitudes/behaviours etc. The implicature results from the markedness of the expression chosen. Consider two examples of markedness implicatures that Nunberg gives:

(3)    *Alas*, the Warriors lost!
(4)    Any junior scholar who stresses teaching at the expense of research *ain't* gonna get tenure.

In (3) the use of the literary and somewhat archaic 'alas' contributes to an ironic tone of the exclamation, which would not be present in 'Damn! the Warriors lost!'. The difference, however, does not follow from the literal meaning of words. An even better example is (4), which Nunberg says is an authentic utterance of a university dean quoted in a *Chronicle of Higher Education* article. Here, with the use of 'ain't' instead of 'isn't', the speaker implies that the truth of what he's saying should be obvious to anyone and does not require an advanced degree ('It ain't rocket science!'). What is important here is that, as Nunberg points out, 'ain't' is not more infrequent or more difficult to process than 'isn't', and it does not belong to a different register (like 'alas'). It is a marked expression in this use because it is governed by the conventions of a (stereotypical) group—say, working-class speakers of English—to which the speaker does not belong.[2] That is how the 'ventriloquistic' implicature that even a 'regular Joe' would accept the observation is generated.

Ventriloquisitc implicatures are generated by a speaker's choice of a linguistic convention that is associated with a different context than the actual one—especially, like in (4), with a different sociolect than the one normally used by the speaker (hence the 'ventriloquism' of speaking as if one was someone else). Nunberg attributes some characteristic properties to these implicatures. First, they are very difficult to cancel.[3] Second, they are always speaker-oriented, even in embedded contexts. Third, they are typically triggered only when the expression chosen by the speaker replaces one that would be the contextual default—that is where the markedness comes from, and without it there is no implicature. In other words, they require the possibility of using a neutral alternative.

These three properties are also the essential properties of slurs: derogatory meanings are non-cancellable, speaker-oriented and they contrast with neutral and respectful ways of referring to the same individuals and groups. Accordingly, Nunberg claims that the derogatory meanings of slurs are (routinized) ventriloquistic implicatures. To support this claim, he assumes that there are different linguistic conventions in different parts of society concerning the *routine* ways of referring to groups. This is prima facie plausible: slurs, at least the common ones, can easily be

---

[2]That the use of 'ain't' implies a low level of education and therefore a working-class background is a stereotype that need not be true—but it is the stereotype that fuels the implicature; for simplicity's sake, I will stick with it in the following discussion.

[3]Note that non-cancellability is not a typical property of conversational implicatures.

identified by a competent speaker as the kind of words that bigots (racists, homophobes etc.) use. And so, by using a racial slur, a speaker aligns herself with the racists and their linguistic conventions.[4] This is, as was already noted, an important function of slurs—to bolster the identity of bigots (as ones that are different from, opposed to slurs' targets).

A useful analogy suggested by Nunberg is to think of the meanings of slurs as exhibiting the same structure a standard dictionary entry would have, identifying the target of the slur with a descriptive term (a neutral counterpart) and adding a metadata label such as 'derogatory':

(5)  **S\***: *derog.* S

Thus, the derogatory force of slurs is a *meta-linguistic* phenomenon. The abstract property 'derogatory' takes on a substantial meaning for any concrete slur through its association with the linguistic conventions of bigots.

As I mentioned at the beginning of this section, the central tenet of Nunberg's theory of slurs is the observation that the linguistics of slurs cannot be fully separated from their politics—and it is in this respect that it differs most significantly from semantic theories. The reason that such and such particular lexical items are slurs is because there are people prone to slur use—and they are not random people nor is their slur use an individual affliction; they exist as part of a broader context of bigoted attitudes, practices and institutions. If not for this context, slurs could not be any different than general insults—'motherfucker', while a vicious insult, is not a slur, because there aren't practices and institutions of discrimination towards a group of people routinely referred to as 'motherfuckers'. Note that this may also explain why slurs targeted at privileged groups (e.g. 'honky'—a derogatory term for white people; or 'wasp'—White Anglo-Saxon Protestant) are rare and rather 'weak' in their offensive capacity—their derogatory power can feed off negative attitudes of their users towards the targets, but has no support in systematic practices of hostility and discrimination. No one really cares much if they're called a 'honky', because it doesn't much matter.

Furthermore, the pragmatic theory easily explains reclaimed uses of slurs,[5] both regarding still offensive words being used by members of the target group in an ironic, subversive, or familiar manner (like the N-word among African-American speakers) and those that have entirely or almost entirely lost their derogatory power, like 'gay'. With regard to the latter, we can simply assume that in the course of a language's history, the slur ceased to be used by bigots often enough to be easily recognizable and available as a trigger of ventriloquistic implicature (possibly because it was replaced by other derogatory words for the same targets). Regarding

---

[4]Cf. Mercier (1995), who suggests that historically, in the USA, using the word 'nigger' amounted to a kind of dialect switch, 'speaking like a Southerner'. What made 'nigger' offensive (rather than an innocuous phonetic variant of the—at the time—inoffensive 'Negro') is that for this reason it was the variant preferred by racists as an expression of their racist attitudes. 'The South's notorious racism thus figures in the lexical history of 'nigger'.' (p. 240)

[5]These are sometimes considered problematic for semantic accounts, see Anderson and Lepore (2013a). On reclamation of slurs generally see Brontsema (2004), Bianchi (2014), Cepollaro (2017), and Ritchie (2017).

the former, in-group users of a slur would prima facie appear as affiliating themselves with the bigots who typically use it against them—which can easily lead to interpretations of their slur use as ironic or subversive. And then again, if this subversive use becomes frequent enough, the slur may finally become neutralized and fully reclaimed—because it has now become the in-group convention to use it; no ventriloquism is involved anymore.

Finally, it should be emphasized that Nunberg's pragmatic theory faces no problem explaining what is perhaps the most important linguistic property of pejorative meaning of slurs:[6] its propensity to project from all embeddings such as attitude reports, conditionals and negation. Ventriloquistic implicatures are by definition not content based and therefore should not be displaced by any syntactic or semantic transformations. Thus, the pejorative meaning of slurs by default projects and does not interact with main, at-issue content.

## 3 Problems with Nunberg's Theory

Nunberg offers a compelling and illuminating account of the nature of slurs and their functioning in social contexts. Nevertheless, the mechanics of his theory are problematic. In this section, I discuss two important weaknesses of the V-implicature account. In the following two sections, I will develop an alternative explanation, based on expressive commitments rather than implicatures.

First, there is a certain incoherence in the way Nunberg explicates what it is that a slur user conveys through their ventriloquistic implicatures and how they do so. As V-implicatures are defined by Nunberg, they presuppose different linguistic conventions associated with different groups of people. Simplifying only a little, the picture is like this: homophobes say 'faggot', non-homophobes say 'gay'. And so, by using 'faggot', a speaker is referring to a convention accepted by homophobes, and thereby implicating that he or she shares the attitudes commonly associated with the use of this slur. The problem here is with the middle step: surely, that 'faggot' is a word for gay men is a convention shared by all speakers of English,[7] as is the convention that it is a derogatory word. This is a problem for Nunberg, because a V-implicature by definition relies on a convention that is *not* the speaker's. A speaker calling gays 'faggots' is a homophobe—this is not somebody else's language.[8] Nunberg could possibly answer that what is important here is which conventions are *default* in which groups—but that won't solve an even more basic problem. 'Bigots' is not a social group that could share linguistic conventions supporting V-implicatures. There is an idiolect of English associated with working-class speakers that includes the default 'ain't' instead of 'isn't'; there are regional idiolects in any language; there are the

---

[6]See Croom (2011), Camp (2013), Jeshion (2013a), and Bolinger (2017), among others.

[7]Or at least American English—wherever one draws the line, there will be bigots and non-bigots on both sides.

[8]This also raises the question of the self-affiliating role of slurs—once a speaker affiliates themselves with bigots, the V-implicature should lose foundation, thereby muting the derogatory meaning of the speaker's slurs.

idiolects of bankers, teachers or literary scholars—but there isn't an idiolect of racists or anti-Semites. Bigots, unfortunately, are everywhere in society and speak all kinds of idiolects. The only thing they have in common is that they are prone to use bigoted language, but that does not get us anywhere, as defining slur use as characteristic of the convention of those who use slurs is obviously circular.

Moreover, even if it was possible to independently define 'bigots' as a social group and attribute a certain linguistic convention to them, this should entail that slurs belonging to it are the neutral, default expressions for those who subscribe to this convention. This may be sometimes true, but as Nunberg is obviously aware, most slurs are a marked, pointed choice for bigots too (that is why they use them after all).

The problem has the form of a dilemma: either slurs are marked expressions even for the bigots who use them or they are the unmarked, default expressions for bigoted speakers. In the latter case, a slur as used by a bigoted speaker would not trigger a V-implicature and therefore we would lack an explanation of why it should have a pejorative meaning or why at all the bigoted speaker would use it. In the former case, the markedness of the slur—on Nunberg's account of how ventriloquistic implicatures work—should result from its association with a specific convention, one that is characteristic of bigots. However, it seems doubtful that slurs can be considered characteristic of some specific convention of language use, such as a dialect or sociolect of some kind. Using slurs is not associated with any specific demographic—it is not a regional variant or something characteristic of the language of a certain class, profession, gender, subculture, religious denomination etc.[9] It is also not associated with any specific context, register or style. Nunberg seems to be aware of this, and he writes that besides all the different possibilities just mentioned and excluded, 'the provenance of a lexical convention can correspond simply to the self-conscious social extension of a certain set of beliefs or attitudes, which themselves can sometimes be inferred from the existence of the word itself'. That is probably true as it stands, but it contributes nothing to a ventriloquistic implicature-based explanation of slur's derogatory meaning. More than that, it seems to suggest that the beliefs and attitudes of bigots, which slurs may be taken to express, have an important role to play here—but if an explanation of the pejorative import of slurs cannot do without reference to the beliefs and attitudes of speakers, it is not a compelling alternative to semantic accounts. Finally, even if one granted Nunberg that slur-using in itself may be a linguistic convention of the right sort, we are back to the other horn of the dilemma. For a bigoted speaker in using a slur does not speak 'like someone else', and therefore should not trigger any V-implicatures. Compare this with Nunberg's non-slurring example: the use of 'ain't' instead of 'isn't'. In the dean's utterance in (4) it triggers a ventriloquistic implicature, because it is *not* the convention that is mutually adopted

---

[9]Mercier (see n. 4) may be right about the history of 'nigger', but this is clearly a special case: for many speakers of American English, there may in fact be an association between racism against African-Americans and a specific geographical region, for obvious historical reasons. This association may easily extend to the regional linguistic variant and thus it could support ventriloquistic implicatures of the kind Nunberg describes. However, this does not generalize to other slurs. 'Kike' may well have originated, as folk etymology has it, on Ellis Island, where illiterate Jewish immigrants signed their documents with a circle (Yiddish 'kikel'), not to use a sign resembling a cross. But using 'kike' to refer to a Jewish person is not in any way to speak like an immigration official, or a New Yorker, etc.

by the speaker and audience in this situation. But now take two speakers for whom 'ain't' is the natural option—stereotypical working-class speakers of English—and when they casually converse, this form will not trigger markedness implicatures of any kind. But when two, say, anti-Semites casually refer to their Jewish neighbours as 'kikes', the pejoration is still present. Which strongly suggests that V-implicatures are not the right mechanism to explain it.

To put this in a different (and possibly simpler) way, there's an irresolvable tension between the idea that V-implicatures involve 'speaking like someone else' and the idea that using slurs is how bigots speak. In the mouth of a habitual slur user, a slur remains strongly marked, but it is not someone else's linguistic convention. There is no ventriloquism here, as the racist speaker speaks very much in their own voice. Thus, it seems that V-implicatures are not the right mechanism to account for the markedness, and derogatory meaning, of slurs.

The second problem is perhaps even more serious, and in any case clearer. Nunberg's theory fails to account for Derogatory Autonomy[10]: the capacity of slurs to offend and derogate regardless of the speaker's intentions. Anyone using the slur 'chink' (in a non-reclaimed way) to refer to a Chinese person is guilty of derogating the referent and all Chinese people, even if they do not actually hold any negative attitudes or beliefs about them—and if even they explicitly disavow any pejorative intent ('I have nothing against chinks; some of my best friends are chinks, etc.' is offensive and derogatory even when spoken sincerely.)

Nunberg acknowledges this fact, writing that 'speakers do bear moral responsibility when they manifest an intention to affiliate with the provenance of a slur in the knowledge that it is not the default term for a group, even when they disclaim any derogatory intent and insist that the word itself is not a derogation at all.' As true as this observation is, V-implicatures are not the right linguistic mechanism to support it. For they need to be deliberate—they cannot result from an unintentional or accidental violation of the Maxim of Manner. To compare slur use with a non-slurring example again, imagine that a word that could be a ventriloquistic implicature-trigger merely slips out on a given occasion—say that a person with a working-class background tries not to say 'ain't' instead of 'isn't' when in the company of people he or she considers better educated. If, despite those efforts, an 'ain't' does slip out, it will probably be taken as betraying the social origins of the speaker, but it will not trigger any implicatures. What that speaker will be taken to *mean* by saying, e.g. 'It ain't gonna happen' is just the same as if they said 'It isn't going to happen', with no implicature of the conclusion being obvious to anyone. On the other hand, the dean in Nunberg's example uses 'ain't' deliberately and therefore triggers the implicature. This is manifestly very different from a situation, for instance, in which a self-aware bigot tries not to use any slurs while in 'polite' company, but in a moment of inattention utters something about selling their house 'before the spics moved into the neighborhood'. The slur is no less offensive and derogatory because it only slipped out. Thus, there is an important disanalogy between the meanings of slurs and ventriloquistic implicatures.

---

[10]Cf. Hom (2008), Anderson and Lepore (2013a), and Camp (2013).

Note that Derogatory Autonomy is a property of slurs generally easily accounted for by content-based accounts—if the derogatory meaning of slurs is conventionally encoded (and not conversationally implicated), the speaker cannot simply intend it away. In 'Some of my best friends are chinks', the slur has the same meaning, on a semantic theory, as in any deliberately derogating use. If Nunberg's proposal cannot account for Derogatory Autonomy, it is not an attractive alternative to semantic theories.

More generally speaking, the meanings of slurs on Nunberg's account are a very peculiar kind of conversational implicatures—they are difficult or impossible to cancel and they are intention-independent. This makes them unlike most conversational implicatures, which in itself should cast doubt on whether this is the right approach. However, in the remainder of this paper I will argue that the core insights of Nunberg's analysis can be preserved if the mechanism of V-implicatures is replaced with another one, free of the weaknesses just discussed.

## 4 Expressive Commitments

The pragmatic theory of slurs which I will propose as an alternative to Nunberg's is based on Hess' (2018, 2019) treatment of the meaning and properties of expressives (such as 'damn' or 'bastard') in terms of *expressive commitments* (EC), which are commitments to the appropriateness of a given expression, by default undertaken by a speaker through the use of said expression. This notion of EC is a development of two conceptions of commitments to appropriateness introduced independently in Brandom's (1983, 1994) philosophical theory of linguistic practice and in Harris' (2014) theoretical and experimental work in pragmatics. In this section, I will present the relevant aspects of the EC theory from Hess (2018, 2019). In the next section, I will apply this conception to slurs and explain how ECs can generate derogatory and offensive meanings.

Unlike Nunberg's markedness implicatures, which are part of a Gricean pragmatic framework, expressive commitments are elements of a discourse structure based on Lewis' (1979) 'scorekeeping in a language game' model. One can think of commitment in the relevant sense as an element of the Lewisian *conversational scoreboard*, with the important difference that the scoreboard has separate rubrics for each of the interlocutors.[11] Thus, participants of a conversation keep and update 'score' on (besides common presuppositions, salience of potential referents, standards of precision etc.) what each of them has committed to. Commitment attribution, in other words, is a way of keeping track of participants' conversational moves. Paradigmatically, in interpreting a speaker's assertion that $p$, a hearer updates the score by attributing to the speaker a commitment to $p$. This will then have an effect on what

---

[11]This picture of scorekeeping and commitment attribution is based on Brandom's (1994) development of the Lewisian model. The crucial aspect of attributing commitments to individual speakers (as separate from updates to the common ground, which corresponds to Lewis' shared scoreboard) has been independently introduced by Farkas and Bruce (2010), as well as Morency et al. (2008)—in, respectively, a formal semantics and cognitive pragmatics setting.

further conversational moves are permitted (e.g. the speaker cannot felicitously deny that *p* in her next utterance).

Besides asserted contents, other categories can also be construed as commitments. Consider one of Lewis's original examples: standards of precision. If Tom is taller than Jerry, and a speaker has said that Jerry is tall, she is now committed to a certain standard of precision for 'tall' such that she cannot in the next sentence felicitously deny that Tom is tall, too. Expressive commitments are another element of this kind.[12] While assertoric commitments are commitments to the truth of asserted contents, expressive commitments concern the way contents (of assertions or other speech acts) are expressed—they are commitments to the applicability or appropriateness of a certain term or expression. They may be salient in a given context because there are importantly different ways of referring to a given object or situation, reflecting diverse perspectives or various associations and connotations. A speaker may wish to bring those perspectives and associations into focus, or downplay them—or they may do that inadvertently. Consider the following examples:

(6)  A: The **freedom fighters** succeeded in liberating the village.
     B: Those 'freedom fighters' butchered its inhabitants. They're **terrorists**!
     (Brandom 1994)[13]

(7)  a. John made Mary a **beergarita**.
     b. John made Mary **what he calls a beergarita**.           (Harris 2014)

(8)  A: Do you want a **pop**?
     B: No, but I'd love a **soda**.           (Hess 2019)

In the exchange in (6), speaker A calls a group of militants 'freedom fighters'—a term with clearly positive associations–which corresponds to his perspective on the events in question (note that 'liberate' is also a marked expression). The speaker B disagrees with the assessment of what happens, and calls the militants 'terrorists'—a term with a strong negative load. Both speakers may not disagree about the facts of the situation (the militants took control of the village, they killed many people), but they do differ in their interpretations, and because of that they take different expressions to be appropriate. A *commitment* to the appropriateness of 'freedom fighters' and 'terrorists', respectively, may be attributed to them and recorded on the conversational scoreboard. (These commitments are recorded in the individual columns of the speakers, because they are not shared.)

Observe also the use of scare quotes in B's first sentence in (6) (in speech, this could be marked with special intonation). They serve to distance the speaker from the expression and thus make it possible for B to talk about the same referent that A does, *without* undertaking the same expressive commitment. This is, it would appear, the main function of scare quotes: to manage expressive commitments.[14]

---

[12]The concept of expressive commitment comes from Brandom, although he seems to reserve it for the use of expressions that differ in inferential significance—this seems to exclude such trivial or extra-linguistic connotations as are relevant in (7) or (8) below.

[13]Example modified.

[14]For an elaboration of this idea, see Brandom (1994).

Other constructions can be used to a similar purpose—Harris (2014) discusses what he dubs transparent free relatives (TFR). A TFR is the kind of construction used in (7-b), a non-sortal 'what' introducing an attributive verb ('call', 'think', 'say' etc.) embedding a noun phrase ('what he dubs transparent free relatives' is also an example, of course). Consider the contrast in (7): in the variant (a.) the speaker uses the term 'beergarita' directly, thereby (implicitly) undertaking the commitment to that term being applicable and appropriate. In the variant (b.), however, the speaker—perhaps judging that 'beergarita' is an extremely silly word—goes out of her way not to undertake this expressive commitment and explicitly attributes the use of the term (and therefore the commitment) to John. According to Harris, managing commitments (of the expressive kind) is precisely what TFRs serve to do.

In those two examples, expressive commitments can appear immediately salient: in (6), the terms 'freedom fighters' and 'terrorists' carry heavy loads of assessment and emotion; a speaker using one of them reveals a lot about their political sympathies. In (7) matters are much less grave, but the name 'beergarita' stands out (for many users of English at least) as an unfamiliar expression, perhaps meant as a joke. Whether it is to be taken seriously, what exactly it means and 'does anyone actually say that' may be salient questions. However, we can imagine that (7-a) is said by one bartender to another, and that 'beergarita' is a standard term for a certain type of drink (the ingredients of which are easy to divine)—in such a situation, the expressive commitment to 'beergarita' would be no more important than the commitment to 'village' in (6). There would be no particular reason to register it on the conversational scoreboard.

Example (8) shows the converse situation. For a large number of American English speakers 'pop' is the standard, unmarked generic term for a sweetened carbonated beverage. Among those speakers, A's way of framing his question would be entirely unremarkable. However, for many other speakers (the divide is geographical) 'soda' is the standard, unmarked term with the same meaning. And so B can, jokingly, bring that up, because for her 'pop' is not a default term. A and B's expressive commitments—to the appropriateness of 'pop' or 'soda' respectively—are relevant because of associations that are largely extra-linguistic.

These examples illustrate some of the main properties of expressive commitments. First, they are independent from assertoric commitments. Regardless of who is responsible for the use of 'beergarita' in (7), whether John or the speaker, the at-issue content of the assertion is the same. Similarly, in (6) it is not the interlocutor's assertoric agreements concerning the facts of the situation that divide them, but the way they both express those facts. Nevertheless, this example also shows that while non-assertoric, expressive commitments can be significant and meaningful—it does make a lot of difference whether the militants in question are called 'freedom fighters' or 'terrorists'.

Second, expressive commitments are always by default attributed to the utterer of the expression. If a speaker wishes to distance herself from a certain expression (like B in (6) or the speaker of (7-b)) they need to resort to special devices, such as scare-quotes, or TFRs. (In direct quotations, of course, the expressive commitments are attributed to the quoted speaker.)

Third, *expressive commitments* are ubiquitous—in a trivial sense, a speaker is by default committed to the appropriateness or applicability of every single word she utters. However, expressive commitments are only salient—and therefore need to be registered on the conversational scoreboard—when the use of an expression is a non-trivial choice with respect to the at-issue content that it is supposed to convey. What makes the choice non-trivial, and the commitment salient, may be any of a variety of factors (which may be more or less contextual). In (6) it is a matter of the strong evaluative and emotive connotations of the expressions 'freedom fighters' and 'terrorists'; in (7) it is a result of the (potential) unfamiliarity of the term 'beergarita'. Moreover, through the use of scare quotes and TFRs the speakers in these examples are deliberately foregrounding expressive commitments. In (8) the cause is the—mutually known—association of 'pop' and 'soda' with different regional varieties of English.

It is important to note that ECs are a category of the *conversational* level, i.e. they are attributed to speakers, with respect to the words they use, in the specific circumstances of a linguistic exchange. They are not properties of words taken as lexical items or as types, but of concrete uses. What makes a commitment to the appropriateness of an expression salient may partly depend on the encoded content of the word (as in the case of 'terrorist') or the sociolinguistic conventions of its use (as in the case of regional 'pop' or jargon-like 'beergarita'), but is ultimately determined by the context of the conversation, and especially the interlocutors' knowledge and assumptions concerning each others' goals, beliefs, and linguistic habits.[15] And so, for instance, in a conversation between two speakers for which 'pop' or 'beergarita' is the default expression (because they share the same dialect or the same expertise), the issue of expressive commitment may not arise at all; it need not be registered on the conversational scoreboard as nothing 'extra' is communicated by it. Similarly, even though words such as 'terrorist' or 'freedom fighter' are inherently charged with evaluative and emotive meanings, the significance of their use depends on the discursive context—what really rides on what a group of militants should be called is determined by the purpose and circumstance of conversation: does it concern events remote in time or space or something happening in the interlocutors' own country right now? is it a conversation between private individuals or a public debate between policy makers? etc. Such issues define the actual salience of the speakers' ECs.

Furthermore, it is only through being salient in such ways that expressive commitments may carry additional meanings. There is nothing significant (outside of very unusual contexts) about calling a dog a 'dog', a drink a 'drink', or, in (6), about the fact that speaker A calls the village in question a 'village'. But by endorsing the labels 'freedom fighter' or 'terrorist', respectively, for the militants who took control over the village, and by using other marked expressions like 'liberated' or 'butchered', A and B are manifesting their political and moral judgments. The speaker of (7-a) may be signalling his expertise in mixology, while the speaker of (7-b) may be making fun of John. The use of 'soda' or 'pop' not only signals a speaker's provenance, but, when

---

[15] For this reason, wherever one draws the line between semantics and pragmatics (nothing in my argument depends on any specific definition of this boundary), expressive commitments fall squarely on the pragmatic side.

deliberate as in B's response in (8), may be a way of making a joke, or of expressing the speaker's emotional investment in their particular variety of English. However, this also entails that the conversational meanings carried by ECs involve, to some extent at least, the interlocutors' extra-linguistic (or meta-linguistic) knowledge and competence: understanding the significance of calling someone a 'terrorist' requires some idea of the political context, while understanding why someone uses 'pop' instead of 'soda' to refer to a carbonated drink presupposes some knowledge about the regional variation of American English. Questions of meaning and competence with respect to ECs will be discussed in more detail in the following sections.

## 5 Slurs and Expressive Commitments

The focus of analysis in Hess (2018) and (2019), of which the relevant aspects were summarized in the previous section, is on the perspectival properties of expressives such as 'damn'—not a pertinent issue in the present context. In this section, I show how the EC account can be extrapolated to the case of slurs, and how it can explain the nature of their derogatory meaning as well as the social aspects of their use that Nunberg focuses on, effectively replacing the V-implicature mechanism with a more suitable one. In the following section, I discuss some further issues regarding slurs that my account can illuminate.

First, let me note that ECs can be construed as a more general category that also covers the effects of ventriloquistic implicatures. Consider Nunberg's example of the dean's utterance of (4):

(4)   Any junior scholar who stresses teaching at the expense of research *ain't* gonna get tenure.

The speaker's use of a grammatical form associated stereotypically with uneducated speakers of English is surprising, and in this context makes his expressive commitment to the appropriateness of *ain't* very salient. The scoreboard should therefore register attribution of this commitment to him. It is a significant aspect of the utterance, because by speaking as if *ain't* was an appropriate expression, the speaker suggests that what he is talking about is something that could adequately be spoken in those terms—which, given the associations of *ain't* with low level of education, can be interpreted as implying that the at-issue content of the utterance is something obvious and easily understood. (The effect, then, is the same as on Nunberg's analysis.)

An important difference between slurs and expressions such as *ain't* is that the former always engender highly salient expressive commitments. Neutral, unmarked terms for social groups—based on race, ethnicity, religion, sexual orientation etc.—are part of linguistic competence. One cannot be a fully competent speaker of English without knowing the words 'Jew', 'homosexual' or 'Chinese', and knowing that these are the basic ways of talking about their respective referents. This trivial observation has the important consequence that using a slur such as 'kike', 'faggot' or 'chink' always engenders a highly salient expressive commitment. A speaker always has a basic, natural, easily available alternative, and thus choosing a slur will always stand

out as a marked option. The utterer of a slur can be taken as committed to the slur being an adequate or appropriate expression, and therefore this is always a salient commitment to be registered on the conversational scoreboard.

As I argued in Section 3, one of the problems that Nunberg's V-implicature account faces is that there is an important disanalogy between the meanings of slurs and markedness implicatures. The use of 'ain't' in a context in which it is the default convention (because it is the unmarked form for both speaker and hearer) does not trigger any such implicatures. But the use of a slur even among speakers who use them commonly does convey a pejorative meaning. And so we should assume that slur use is never the default convention. But in that case there is no target for the 'ventriloquism' of Nunberg's implicatures. However, expressive commitments do not require any kind of ventriloquism or a distinction between one convention and another. It is a *shared* convention of speakers of English that 'Korean' is the default term for Koreans, and 'gook' is the derogatory term for Koreans. That is precisely the kind of meta-linguistic information which is, as Nunberg observes, recorded in dictionaries, and it is part of linguistic competence of English speakers. That both terms are easily recognizable to speakers for what they are—a neutral or a derogatory term, respectively—is enough to make an expressive commitment resulting from slur use highly salient.

The salience of an expressive commitment, however, does not yet fully account for the derogatory meaning of a slur. The question that has not been answered so far is, how do expressive commitments give rise to meaning? Which, I propose, can be derived from an answer to the question: what does it mean to be committed? For this, it will be useful to refer to Brandom's (1983, 1994) conception of commitments, both assertoric and expressive. To undertake an assertoric commitment, according to Brandom,[16] is to take on a justificatory responsibility with regard to the content asserted. Thus, a speaker asserting that a group of militants took control of a village, takes responsibility for the truth of this statement. The exact nature of this responsibility is determined by the kinds of challenges that the speaker may now face—in the first place, a justificatory challenge, if anyone should question the truth of what was asserted ('I know it happened, because I read about it in a reliable news report etc.'). For expressive commitments, there is no truth to be challenged or defended, as they are not propositional, but their importance for Brandom lies in the fact that they have some kind of inferential significance: they can provide premises or conclusions of additional inferences. If the same speaker asserts that *the terrorists* took control over the village, her words have extra significance—we're not just talking about militants now, but terrorists. Expressive commitments of this sort also constitute justificatory responsibilities, which determine what the commitment actually means. And so the significance of calling the group in question 'terrorists' rather than something else lies in what the speaker would say if challenged on that—how they would justify the appropriateness of a loaded expression, and what kind of inferences would license the use of this term (as well as what kind of inferences can follow from accepting

---

[16]Note that for Brandom this is an essential aspect of what it is to make an assertion and part of the explanation of what the *content* of an assertion is. For my purposes it is enough to treat it as one of the properties of assertive acts.

that someone is a terrorist). Note that this may involve various reasons—a moral judgment, a political position, a legal argument etc.—each of which could provide the answer to a question 'Why do you call them 'terrorists'?' or 'What do you mean by 'terrorists'?'. In any case, however, these reasons go beyond the strictly semantic content of the expression in question and concern the significance of the speaker's conversational move—the question is not what 'terrorist' means, but what the speaker means by 'terrorist'. And while these two issues are obviously connected, only the latter belongs properly to the level of conversational dynamics (on which I propose to analyse the derogatory meaning of slurs).

Brandom's notion of justificatory responsibility can be adapted to my purposes in the following way, focusing on the ramifications of the speaker's choice of expression rather than on inferential relations between expressed propositions. The meaning of a salient expressive commitment can be construed as whatever would provide the answer to a question 'Why do you use that word?'. The reasons can be of different kinds. In the case of 'terrorists', it may be a moral or political judgment; in the case of 'bastard' a negative attitude towards the person described with that epithet, or a heightened emotional state—in other words, not necessarily something that can be expressed in a proposition that would provide a premise for an inference, but a concrete reason nonetheless. These reasons will simultaneously be what those salient expressive commitments communicate. A speaker may use 'pop' instead of 'soda' out of a (more or less serious) commitment to a certain variety of English, to tease their interlocutor, to signal that they consider their variety to be superior or 'more correct', or they may do so unreflectively, because that is simply the word that is used where they grew up. In the latter case, their use may still be meaningful—precisely in that it reveals something about their identity. (Naturally, in each of these cases the reasons why a speaker takes a certain expression to be appropriate are not typically explicitly given or asked for, but only inferred by hearers.)

On the view I propose here, slurs are in an important respect more like 'soda' and 'pop' than like 'terrorist' or 'bastard'. As Nunberg argues, there is no need to suppose a semantic difference between a slur and its neutral counterpart, just like there need not be a semantic difference between 'soda' and 'pop'. What an expressive commitment undertaken in the use of a slur *means* is entirely a matter of social context, not literal content. However, unlike in the case of terms for carbonated beverages, where the choice of one or the other may be only rarely significant, the commitment to the appropriateness of a slur is always salient and never trivial. For slurs are linked to practices of discrimination and injustice. They are, indeed, tokens of those practices. By undertaking a commitment to the effect that the slur is an appropriate expression, one is *ipso facto* committing oneself to it being appropriate to denigrate and offend the people who are the slur's targets, because that is the load that the slur carries in virtue of its functioning in social contexts of iniquity and violence. A given word is a *slur*, a toxic, offensive expression, because it is part of practices of bigotry. While injustice and discrimination have many different faces, some more sinister than that, verbal abuse is definitely one of them. If knowing that a given word is a derogatory slur is something one can expect from a competent speaker of a language, as it surely is, than one can also expect them to be aware of their role in hateful practices. The latter aspect may not be part of linguistic competence in any strict sense, but it

is undeniably something that belongs to most language users' common and mutual knowledge, and it is what uses of slurs mean in actual, real-life situations.

Three aspects of the EC theory of slurs' meanings require further clarification. First, the meanings conveyed by ECs need not be—and in the case of slurs, they are not—semantically encoded in the expressions; but they are also not (merely) a function of speakers' communicative intentions. They depend essentially on *hearers'* inferences and assumptions concerning meta-linguistic facts. Just like in the case of the ECs discussed by Harris and Hess, to interpret a speaker's use of a salient expression, the hearer must take into account more than just what the speaker believes or intends—they must make assumptions or inferences about whether or not 'beergarita' is a common word, a jargon expression, or a joke; who and why uses 'pop' rather than 'soda' etc. With regard to slurs, this involves the aspects just discussed: that slurs are commonly known to be derogatory words and that they are employed in practices of bigotry, discrimination and violence. It is based on the hearer's awareness of these social facts that slurs mean what they do. This also implies that hearers may differ regarding their interpretation of a given expression and even in their judgment of whether it is a slur or not—as I will discuss in the next section. More importantly, the meanings that hearers attribute to slurs on any given occasion of use may be much less specific than any attitude, emotion, belief or stereotype that an individual speaker holds.[17] On the account I propose, the essential aspect is that a slur user undertakes the commitment to the appropriateness of an expression that is commonly known to be derogatory, because it is employed in bigoted practices—the derogatoriness and offensiveness of the slur is inherited from those practices, which are multifarious and complex. There is no need to assume that every user of a given slur is communicating the same specific content, or even holds the same hateful or contemptful attitude towards the target group.

The second consequence of this is that slur uses may convey derogation regardless of the user's attitudes and intentions. If we can assume that a speaker is aware of the functioning of slurs within practices of bigotry (I will discuss less-than-knowledgeable speakers in the next section), and yet he or she chooses to use a slur, thereby undertaking a commitment to this expression being appropriate, we can also take the speaker to accept those associations. Even if the speaker does not fully endorse the bigoted linguistic and extra-linguistic practices that the slur is associated with, or if he or she has no active intention to offend or denigrate, the expressive commitment is sufficient to engender derogation, because those associations are a social reality, independent of the speaker's beliefs and intentions. Choosing a slur as the adequate expression is already derogatory, because it entails the acceptance (deliberate or not) of a practice of bigotry which is manifested, among other things, in the use of derogatory language. That is, a speaker using a slur manifests, at the very least and almost tautologically, that they do not mind being interpreted as participating in a practice of slur use. This does not have to mean they positively endorse it; it's enough that they go along with it (it also does not mean that the speaker

---

[17]Cf. Jeshion ([2013b](#)) who argues that slurs cannot be associated with any specific stereotypes, because slur users can deny any and all stereotypes while still successfully derogating their targets.

belongs to a 'bigoted community'; practices of bigotry are, sadly, not limited to members of any specific communities, but ubiquitous throughout the society). Because the existence of this practice, and the slur's association with it, is not dependent on the speaker's intention, neither is the derogation conveyed. Thus, the EC account, in contrast to Nunberg's, has no trouble explaining Derogatory Autonomy.[18]

Third, furthermore, understanding the derogatory meaning of slurs is a matter of grasping the broader social context of their typical use, and not merely their encoded content. Which in turn requires some awareness of the practices of bigotry and discrimination of which the slurs are tokens. In other words, understanding slurs requires more than strictly linguistic competence.[19] And this holds not only for the hearer's side: speakers are generally taken to be responsible for the derogation effected by their slur uses, even if not intentional, because it can be assumed that, as competent members of a society, they are aware of the embedding of slurs in practices of bigotry. In other words, interpreting the significance of expressive commitments undertaken by the use of slurs requires both hearer knowledge of the extra-linguistic context and an assumption of speaker knowledge of that context. In Section 6.2, I discuss situations in which this assumption is not in force.

To recapitulate my proposal so far, the EC theory offers an important improvement over Nunberg's implicature-based account, while preserving the general picture of the 'social life of slurs'. A speaker using a slur undertakes a highly salient expressive commitment (i.e. the hearers may attribute such a commitment to the speaker, and it is registered on the conversational scoreboard). The meaning of this commitment can be inferred by the audience based on assumptions about both the speaker's beliefs and attitudes, and, crucially, about linguistic and meta-linguistic facts, especially those concerning the functioning of slurs within practices of bigotry. This account faces no problems with defining 'speaking like a racist' in terms of distinct default conventions, which is an important weakness of Nunberg's theory, as I argued in Section 3. Speakers prone to slur use and those that treat it as objectionable share the same conventions, but they engage in different practices. Another important advantage my account offers when compared to Nunberg's is a straightforward explanation of Derogatory Autonomy. Unlike V-implicatures, expressive commitments are entirely intention-independent, and so are the associations that link slurs to bigoted linguistic and extra-linguistic practices. These can provide grounds for attributing a derogatory meaning to a slurring utterance, even in the absence of a derogatory intention. There are ways of avoiding expressive commitment when uttering a certain word

---

[18]Jeshion (2013a) presents a similar view on Derogatory Autonomy, arguing that hearers may be justified in taking the use of a slur to be derogatory (to express a derogatory attitude, on her account) even if the speaker has no such intention.

[19]Cf. Vallée (2014, p. 86): 'Grasping the cluster of negative ideas a slur conveys in an utterance requires non-linguistic, ordinary knowledge of the underlying prejudices against a community.' Note that this is entailed even by some strictly semantic accounts, e.g. Jeshion (2013a) argues that there need not be a semantic difference between slurs and other pejoratives, without denying that the offensiveness of slurs is unique; to understand it, then, one has to understand the social context of slur use.

(quotational and distancing devices, such as scare quotes or TFRs[20]) but in the absence of them, a speaker is committed to the adequacy and appropriateness of the words he or she actually uses—and those words are what they are, whether one likes that or not. Commitments are by definition public and objective.

Finally, it should be noted that the EC theory can also explain the projective behaviour of slurs (in a way essentially similar to Nunberg's): because the derogatory meaning is not part of the semantic content of a slur (which is identical to its neutral counterpart) it cannot be a target of semantic displacement due to negation, temporal or modal operators, embedding in speech reports and conditionals etc. More specifically, the offensiveness and derogation carried by a slur are the function of a speaker's choice to use that expression instead of an available neutral alternative. Outside of quotational contexts (broadly construed, including scare quotes and TFRs), the choice of words is always the full responsibility of the actual speaker, whatever the syntactic and semantic complexity of an utterance.[21]

In the next section, I discuss several further issues which can be illuminated with help of the EC theory of slurs (including a brief remark on reclaimed uses).

## 6 Further Applications

### 6.1 Derogation and Offensiveness

Slurs derogate and offend the groups and individuals that are their targets. But slurs also offend in another way: using them is a transgression, and it is offensive to any (non-bigoted) hearers, even if they are not themselves members of the target group. The offense to targets and to hearers are of course of very different kinds, but they are both characteristic effects of slur use. Not being Jewish, I am not directly harmed by anyone using the slur 'kike'—I am not the target of the hateful practice it is a token of nor am I humiliated or insulted by it—but nonetheless I am rightly offended by it.

The account I propose here makes it possible to explicitly distinguish those two aspects (which is rarely done in the literature on slurs). I will reserve the term 'offensiveness' for the effect of slurs on (non-target) hearers, and 'derogation' for their pejorative meaning which concerns targets. A distinction of this kind has been previously proposed by Hom and May (2013; cf. Hom 2008), although in a different setting and for a different purpose. Hom and May construe the derogatory meaning of slurs in truth-conditional terms, which raises the question why slurs are still offensive even in negated contexts ('Yao is not a chink' seems no less offensive than 'Yao is a chink'). Accordingly, they propose to distinguish offensiveness (or 'offense') as a psychological category from derogation as a semantic one. The latter has to do

---

[20]Meta-linguistic negation is another important way of disavowing commitment. Consider an utterance like 'I'm not a kike, I'm a Jew.'; it is clear that the speaker does not actually *use* the slur and does not undertake a commitment to its appropriateness—quite the contrary. On meta-linguistic uses of slurs, see Anderson and Lepore (2013a) and Whiting (2013).

[21]For this reason slurring meanings cannot be displaced by embedding or targeted by direct negation—but they are suitable targets of meta-linguistic negation, as noted e.g. by Whiting (2013).

with the literal content of a slur, which is subject to negation. However, offensiveness persists, but for Hom and May it is not an aspect of the encoded meaning of slurs. The main rationale for this distinction is to explain away the appearance of projective behaviour of slurs. In an utterance of 'Yao is not a chink', on this view, the derogatory meaning of the slur is negated, but the utterance may still be considered offensive, e.g. because it conversationally implicates that the speaker does not consider the predicate 'chink' to be empty, but only refrains from applying it in this particular case. This, Hom and May argue, is not a function of the meaning of the slur *per se*, but an effect that depends on the hearer's psychological states, the formality of the context etc.

As argued by Sennet and Copp (2015) and Cepollaro and Thommen (2019), Hom and May's strategy of dealing with apparently projective behaviour of slurs is not a successful one. It seems much more plausible that it is in fact an essential property of the meaning of slurs, as almost all other theorists assume. Therefore, a distinction between derogation and offensiveness along the exact lines Hom and May draw it is not useful. Moreover, it does not capture the aspect I am interested in here: the different kinds of offense to targets and hearers of a slur. Hom and May's characterization of offensiveness as a psychological effect appears to be limited to offensiveness to the actual audience of a slurring utterance, which may include both targets and hearers. What they do not seem to take into account is situations in which no member of the target group is present. If offensiveness in the case of negated slurs is just a psychological phenomenon, it seems that utterances like 'Yao is not a chink' would not offend (or derogate) Chinese people, if no Chinese person hears them—and that is a counterintuitive and troubling consequence.

This is not to deny the reality of the psychological effects of offensive language; I merely claim that they cannot be used to fully account for the projective behaviour of slurs. An utterance may cause offense for a variety of reasons that need not have anything to do with the strictly linguistic properties of the words used by the speaker. 'Jewish people are by nature clever and greedy' is an offensive statement, but not in virtue of its linguistic features—consider the fact that a report of the form 'Kiki said that Jewish people are etc.' does not in itself offend Jewish people. This kind of offense does not project from embeddings, and therefore it is a phenomenon distinct from what we usually have in mind when we talk about the offensiveness of slurs (although it may concern slurs as well). I will not have anything more to say about it.

It is not my aim here to criticize Hom and May's account in detail, but to offer a different way of conceptualizing derogation and offensiveness of slurs. First of all, offensiveness should not be construed as merely an extra-linguistic byproduct of slur uses, as it is an essential property of slurs. Support for this can be found in Bolinger's (2017) distinction between *actual* and *warranted* offense.[22] The use of a derogatory term warrants offense (because of its derogatory meaning) even if no

---

[22]Bolinger further distinguishes situations in which offense is *rational*, even though it is not strictly warranted. One such situation is when a speaker uses a slur without knowing it is a derogatory word—I discuss this kind of situation in the next subsection.

one actually *takes* offense at it. A slur may go unnoticed or the audience may agree with the speaker's derogation of its referent—this should not entail that a slur is not actually offensive. Hearers *should* take offense at speakers' (non-reclaimed) slur uses, because they irrevocably offend their targets. Bolinger's argument combines both aspects that I propose to distinguish here: derogation of targets and offensiveness to hearers, but on my view both should be construed as warranted rather than merely actual. That is, slurs should be taken as derogating their targets and as offensive to hearers regardless of whether anyone actually feels derogated or offended on any given occasion of use.

With this in mind, I propose to understand offensiveness of slurs (in a non-psychological or reductive sense) along the lines of Anderson and Lepore's (2013a, b) Prohibitionism. Anderson and Lepore propose to understand the offensiveness of slurs as a function of their taboo status: there are social norms in place that prohibit the use of slurring terms (hence they call this theory Prohibitionism); any utterance of a slur is a violation of such a norm, which is the source of its offensiveness. Following Prohibitionism, we can say that offensiveness results from a violation of a norm that prohibits speakers from uttering slurs. Derogation, however, is a function of the bigoted stance which a hearer is justified in attributing to a slurring speaker. It does not result from a mere norm violation, but from what this violation represents. Anderson and Lepore's explanation should be limited to offensiveness in the narrower sense which concerns hearers, and not targets. Moreover, it seems only natural to assume (as Anderson and Lepore are forced to deny) that derogation of the targets, which is a distinct aspect, is the reason why there are norms prohibiting the use of slurs. Slurs are offensive (and prohibited) because they are derogatory.[23] Finally, on my view, it is not the mere utterance of a slur, but the commitment which it creates that is the source of offensiveness and derogation. Uses that are non-committal (e.g. quotative uses) may therefore avoid being offensive and derogatory.

Derogation and offensiveness can be distinguished as resulting from two distinct aspects of the interpretation of slur uses in the model described in the previous sections. A speaker using a slur undertakes a commitment to its appropriateness. This is offensive to any hearers, because slurs are essentially inappropriate. They share this property with swear words and other taboo expressions, although their inappropriateness is typically much more toxic. A speaker who uses such an expression and therefore presents it as acceptable and adequate, violates a norm and commits an offense. Derogation, on the other hand, is a function of the slurs' status as a token of bigoted practices. It is what essentially distinguishes slurs from swear words. Saying 'fuck' at the family Christmas dinner is offensive, but it does not derogate anyone. Saying 'faggot' is similarly offensive (or more so), but it is also an act of a practice of bigotry and derogates gay people. Like swear words, slurs are subject to a norm saying they are words that *should not be used* (the scope and strength of these respective

---

[23]In general. See below for a discussion of instances where a given tokening of a slur is offensive, but not derogatory—it is however offensive only for the reason that *normally* it is derogatory.

norms, as well as their sources, may be very different, but they are similar in their prohibitory nature), but there is a great difference in what they are used *for*.[24]

To make this distinction clearer, we may imagine a hypothetical language community, in which everyone is so oblivious and complacent with regard to practices of racism and discrimination that they are insensitive to the derogatory aspect of slurs; they just habitually use them to refer to groups they hold contempt against, and see nothing especially wrong with it. However, the words are still considered by these speakers to be slang or cuss words, rude expressions that should not be used in polite company. For members of this community slurs would be offensive (to hearers) but not derogatory with respect to their targets.[25] They would be like the word 'fucking' used to refer to sex in actual English—an expression that does not carry any pejorative associations whatsoever, but one best not uttered during, say, a job interview. By contrast, a word that used to be a slur but ceases to be one, like 'gay' (in most contemporary uses, it would appear), is neither derogatory nor offensive to hearers. It does not give rise to a salient expressive commitment anymore than 'straight' does.

The distinction between derogation and offensiveness captures an important aspect of the functioning of slurs; in one way they are similar to swear words and expletives, but in another they are very different.[26] It will also be useful in the discussion of further aspects in the next two subsections.

## 6.2 Less-Than-Knowledgeable Speakers

The EC theory of slurs, and the distinction between derogation and offensiveness it supports, makes it possible to account for conflicting intuitions concerning an interesting, but rarely discussed issue: slur-utterances by ignorant or less-than-knowledgeable speakers (e.g. children or foreign language learners). On the one hand, it seems we ought to be able to exculpate them, at least partially, for uttering a slur the meaning and status of which they are not aware of. On the other hand, we usually feel that their slur uses are still in some way offensive (imagine a 5-year-old running around calling Black people 'spooks', just because he thinks it means nothing more than a dark figure). This may pose a difficulty for other theories of slurs. On semantic accounts, we should expect derogation to be fully present in such situations, because the pejorative meaning is part of the semantic content. If offensiveness is a function of that meaning, it should also be fully present, and it is difficult to see why a speaker could be in any way excused. Prohibitionists would also have to agree that those uses of slurs are equally offensive (in their broad understanding of

---

[24] As suggested by an anonymous reviewer, a further distinction may be made here: the offensiveness of swear words is only a matter of offending audiences through a breach of linguistic decorum, while the offensiveness (to audiences) of slurs is derivative of their derogation (of targets). While I agree with this observation, I do not believe it to be clear whether this is a difference in the kind of offense or difference in the kind of normative prohibition (impoliteness vs. bigotry) the breaking of which gives rise to offense. The matter requires further investigation, ideally one informed by empirical research on audience response to different categories of pejoratives.

[25] On the status of racist slurs as slang words—thus, ones that may be inappropriate in this weaker sense even for those who endorse their use—see Hill (2008), also discussed by Nunberg.

[26] On the difference between slurs and expletives, see also Blakemore (2015).

offensiveness concerning both hearers and targets of slurs), because a norm is being broken all the same, whether the speaker is aware of it or not. On Nunberg's view, on the contrary, there should be no problem with such utterances at all—clearly, no ventriloquistic implicatures are generated, if the speakers are not aware of the differing conventions for the words they use.

Bolinger (2017), who addresses the issue directly (although very briefly), comes closest to a satisfying solution, by claiming that speaker ignorance can undercut the link between the use of a slur and the endorsement of associated attitudes that this use normally signals (in competent speakers). It remains, unclear, however, if and why the ignorant use should be problematic in any way if this link is absent. Bolinger claims that in such case offense taken by hearers is rational, even though it is not strictly speaking warranted, but that is not a clear distinction. We should rather construe such slur uses as offensive (in the narrower sense I proposed in the previous subsection), but not derogatory. Whoever utters a slur in a non-quotational context undertakes a commitment to the effect that the slur is an appropriate expression (i.e. hearers are warranted in attributing such a commitment to the speaker). This is ground enough for taking offense—slurs are never appropriate expressions, and hearers should feel obliged to sanction and protest their use. However, uses of slurs by incompetent or unknowing speakers need not be interpreted as derogatory, because those speakers are unaware of the functioning of these words in practices of bigotry and discrimination—and as I have argued in this essay, it is those practices that imbue slurs with their derogatory force. A small child saying 'spook' does cause offense, and should be corrected, but she is clearly not participating in a practice of racism. The same may go for language learners; both categories of speakers may have not only a limited linguistic competence but also a limited *social* competence.[27] Just as they could be excused for violating some norms of politeness or civility—it's possible for them to do something that is frowned upon, like addressing another person inappropriately, without actually *being rude*—their utterances of slurs need not carry derogatory force.

In other words, the use of a slur by an unknowing speaker is offensive, because a slur is a word that should not be used; and it should not be used because of the hateful practices in which it is involved. That is reason enough to take offense to and sanction all uses of slurs.[28] However, derogation requires an additional element, which is that the speaker understands what they are doing in uttering a slur—and a less-than-competent language user may not understand this and thus may be excused. This, of course, is largely for the hearers to decide, because they need to make assumptions and inferences about what the speaker is aware of and can be taken responsible for. Note, furthermore, that lack of knowledge is a very different situation from lack of intention to derogate. A speaker who is aware of the status of a slur may be taken to task for its derogatory meaning, even if they try to disavow it; but a speaker who is

---

[27] As noted in Section 4, a full appreciation of the derogatory meaning of slurs requires a linguistic competence that goes beyond the mere grasp of encoded contents. This should not appear problematic, as many things we expect of competent language speakers require more than semantic knowledge: understanding the use of formal and informal terms of address, nicknames and honorifics, polite and rude expressions etc.

[28] As opposed to mere mentions in quotational contexts, which may avoid being offensive at all.

not aware of it may be excused (although non-bigoted hearers would typically, and rightly, feel that they should correct them).

## 6.3 When Is a Word a (Derogatory) Slur?

Commitments are properties of speakers, not of utterances (unlike semantic contents). On the other hand, in contrast to ventriloquistic implicatures, they are intention-independent. This makes the concept of commitment well-suited to account for the unique entanglement of linguistic and socio-political matters which can be observed in slurs. For commitments are attributed to individuals—concrete persons in a social reality. Their significance can therefore be sensitive to extra-linguistic context. Moreover, this entails that the derogatory meaning of a slur is not an object of merely *linguistic* competence in some narrow sense. And so, for instance, understanding what 'homosexual' means may be just a matter of linguistic competence, but understanding the full meaning of 'faggot' requires at least some awareness of the bigoted practices of which this word is part, whether the speaker condones or rejects such practices. (Hence, as argued in the previous section, speakers who are not aware of this social embedding of a slurring term may be partly excused for using it without proper understanding.) It also entails that the status of slurs as derogatory terms may be subject to interpretation, differences of opinion and controversy in ways in which the meanings of 'ordinary' words are not. If derogatory meanings depend on the embedding of slurs in social practices, it is possible that those practices may be themselves variously interpreted and assessed by language users. This has consequences for the understanding of reclaimed uses, as well as controversies concerning the status of certain words as slurs.

Reclamation, or reappropriation, of slurs is a complex and difficult topic.[29] A full account of how the expressive commitment-based theory helps to explain it is left for future work, and here I will make only some initial remarks. The conceptual puzzle of reclaimed uses of slurs concerns so-called in-group uses,[30] i.e. utterances of words that are generally considered slurs by members of the respective target group with a non-derogatory, often ironic or subversive intention (e.g. 'nigger'/'nigga' as used by African-Americans, especially in artistic and comedic contexts, or 'dyke' as used by gay women, especially in political activism). The difficulty in explaining in-group uses is that they demonstrate a sensitivity of slurs' offensiveness and derogation to the social identity of the speaker. Imagine a Black comedian (say, Dave Chappelle) presenting a routine that involves many utterances of 'nigger' to refer to fellow Blacks. For a sympathetic audience, his uses of the slur will not appear derogatory (even if they retain some residue of offensiveness, because it is still the same word, that *generally* should not be used). That is, the utterer of those slurs might not be taken to be derogating and insulting African-Americans; his use may be acceptable. An innocent

---

[29] See the references mentioned in note 5.

[30] To be distinguished from the case of words that used to have derogatory meaning, but became neutralized as a matter of historical development, e.g. 'Mormon' or (to a large extent) 'gay'. This is an interesting phenomenon, but not a very puzzling one: words generally change meanings over time; and moreover, non-derogatory uses of 'Mormon' or 'gay' are not limited to members of the target groups.

interpretation of his speech can be facilitated by the context of a comedy performance and a manifest lack of derogatory intention, but it has been argued already that lack of an offending intention is not nearly sufficient to remove derogation. And neither is the context alone—which is easy to see if one only imagines that a white comedian would try to do the same[31]; even if a particular monologue or sketch was repeated word for word, with manifestly the same intention and attitude, the effects would be hugely different. Non-quotative uses of slurs by out-group individuals never fail to be derogatory. It is the social identity of the speaker *as an in-group speaker*, i.e. as one of the usual targets of the slur, that enables a different interpretation of the use of a slur. It is race in this case, but there are analogous situations for other aspects of identity and in-group uses of slurs that target those aspects; the contexts need not be only comedic. The word 'dyke' has a very different significance when used by a random straight man than when used in the name of the LGBT-activist organization 'Dykes on Bikes'.

In-group uses can be accommodated within my account of slurs in virtue of the fact that the direct discourse effect of a slur-utterance by a speaker is the attribution of the respective EC to that speaker. He or she now counts as committed to this word being adequate. In normal circumstances, this commitment entails that the speaker accepts (whether willingly or not) the practice in which this slur is embedded— the practice of bigotry and contempt, manifested among other things in hostile and derogatory speech. It is with such practices that slurs—as marked alternatives to easily available neutral and respectful ways of talking about people—are associated, to mutual knowledge of competent language users. Those associations cannot be shaken off, even if the speaker sincerely intends no harm or insult to anyone, but they can be *reinterpreted* in the context of *alternative practices*, such as comedy or political activism. In those contexts, the same associations of bigotry and contempt have a different status—they are now targets of mocking or resistance, rather than objects of acceptance (whether enthusiastic or complacent).[32] That is where the importance of the fact that commitments are properties of speakers lies: the social identity of a speaker is something that may enable her or him to participate in a practice that is unavailable for others. And so, for instance, the practice of using 'nigga' to refer to Blacks as an act of subversion or irony, or in-group solidarity, is one that only Blacks can participate in. It's not that others, whites in particular, are not 'allowed' to (as racists often complain)—it's just a different practice when they utter the same words, and so the effects are different too. Note that the fact that one must meet certain conditions to participate in a social practice entails that whether or not a given use of a

---

[31] When Michael Richards, a white man, in an infamous incident during a comedy performance in 2006, addressed a 'nigger'-filled rant to an audience member, it caused immediate and universal outrage; he was not able to explain his slur use away as part of the comedy (although he was clearly tempted to do so, as later interviews show). See 'Seinfeld actor lets fly with a racist tirade', *The Guardian*, 22 November 2006 (www.theguardian.com/world/2006/nov/22/usa.danglaister—accessed on April 24, 2019).

[32] Practices which allow non-derogatory uses of slurs may be less definite and robust than this—friendly teasing may be one, but not everything goes. A Black teacher cannot address his Black students with 'niggers' nor can a lesbian politician refer to her wife as 'dyke' during state dinner, even if she may be able to do that at a rally for LGBT rights.

slur can count as reclaimed and non-derogatory does not depend only on speaker's intentions—as an analogy, consider the practice of democratic elections; someone's filling out a ballot and placing it in an urn may count as an act of voting or as election fraud depending on whether or not they are a properly registered voter.[33]

However, the problematic associations of the slur are reinterpreted, but not removed—hence the remaining 'sting', a residue of offensiveness (which is so valuable for comedy), but also the potential for controversy, as not all recipients are equally willing to accept the reinterpretation.[34] This last aspect can also be explained on the EC account: the actual significance of an expressive commitment is partially dependent on the assumptions and inferences that hearers make in attributing it to the speaker. And hearers may differ in what they are ready to assume: in particular, they may differ in their opinions of the sense and value of the 'alternative practices' of in-group slur use—and if they do not condone those practices, they may consider those slur uses to still carry their full derogatory meaning.

The status of a word as a supposed slur may be contested on a different plane as well. As Nunberg points out, 'slur' itself is a thick term, combining description with evaluation. Due to the presence of an evaluative component, as with any other thick term, individual judgments concerning the application of 'slur' can differ. People may disagree whether a given word is a slur or not. A high-profile example from a few years ago is 'Oriental' as used in official and quasi-official language in the United States for individuals of Asian descent.[35] A few years ago it was removed from the US federal administration's documents, because it was judged to be offensive (a slur in the sense in which this terms is used in philosophical debates). The decision was welcomed by many who felt that the word was indeed disrespectful. Others, however, argued that there was nothing wrong with it and that it was a perfectly neutral, if somewhat archaic, term for individuals of Asian provenance. A similar debate surrounded the word 'allochtoon' (a person with immigrant background) in the Netherlands; the government has also stopped using it.[36] Even more strikingly, speakers can differ about a given word's status as a slur, even if everyone agrees that it is a pejorative. This is especially true of expressions targeting aspects of person's identity other than ethnicity, race, religion or sexual orientation—ones referring to political views or professions. 'Fascist', as used to refer to nationalists and radical right-wingers, is sometimes argued to be a harmful and unacceptable slur (especially by those targeted by it), while others consider it an adequate (pejorative) term for actual fascists. 'Lefty' or 'socialist', on the other hand, seem to be judged as more harmful by those who use them with the intention to derogate than those who are

---

[33] An account of reclaimed slurs based on such a conception of social practices is developed in detail in Hess (2020).

[34] For the case of 'nigger'/'nigga', see Kennedy (2003) and Asim (2007).

[35] Cf. 'The term 'Oriental' is outdated, but is it racist?', *Los Angeles Times*, 1 June 2016 (http://www.latimes.com/opinion/op-ed/la-oe-tsuchiyama-oriental-insult-20160601-snap-story.html—accessed on April 24, 2019).

[36] Cf. 'Overheid schrapt 'allochtoon' per direct uit vocabulaire', *De Volkskrant*, 1 November 2016 (https://www.volkskrant.nl/binnenland/overheid-schrapt-allochtoon-per-direct-uit-vocabulaire~a4406224/—accessed on April 24, 2019).

actually their targets ('lefty' is plausibly a case of reappropriation). Words like 'shyster', 'whore' or 'junkie' are also subject to controversy—they clearly convey a negative attitude towards their referents, and many of them may be considered taboo swear words unacceptable in polite conversation, but opinions may differ on whether or not they are offensive and unacceptable in the same way as racist or homophobic slurs.

Debates, controversies and doubts of this sort can be naturally accounted for in my model, because it relies on the attribution of expressive commitments and associations with bigoted linguistic practices to speakers by hearers. Hearers with different background beliefs and attitudes may interpret a given use of an expression differently, because they may connect it to different kinds of practices, or they may interpret and assess the relevant practices in various ways. For some users of American English, 'Oriental' will invoke associations with practices of linguistic pejoration and extra-linguistic discrimination of Asian-Americans which they deem reprehensible; for others those associations may not be salient, perhaps because they are entirely unaware of them (the word may not be marked for them, so the issue of the speaker's expressive commitment never even arises). In the case of words such as 'fascist', those targeted by them may perhaps feel that they are unjustly denigrated for their conservative political views and that 'fascist' is a slur invoking and reinforcing this injustice. Others may consider the practice of calling out fascists just and proper, because they do not think that the targets' political positions deserve any respect. Similarly, 'whore' is presumably not a slur for those who consider sex workers contemptible because of their life choices and therefore find nothing morally wrong in the practice of calling them names. In each of those cases, a speaker using the word in question undertakes the expressive commitment to this word being an appropriate label for its respective referent. The significance of that commitment, however, is left for the hearers to interpret. Depending on what associations they have linking the word to wider linguistic practice, they may deem the word a slur or not. This is not to say that it is a matter of individual liking whether some expression is a slur or not. The factors involved are objective and intention-independent. But they are social and political, and as such they are always fragmented and constantly contested.

## 7 Conclusion

Slurs are a both troubling and intriguing category. Their meanings are sensitive to social divisions and conflicts in a unique way. Geoffrey Nunberg proposed a theory that offers substantial insight into their meaning and nature. However, as I have argued, the exact pragmatic mechanism he stipulates is not up to the task. Accordingly, I developed an alternative account, based on the concept of expressive commitments, that covers the same explanatory ground as the ventriloquistic implicature theory, while avoiding its problems. Moreover, the expressive commitment model of slurring meanings supports accounts of little discussed, but important aspects of slur use, including a distinction between derogation and offensiveness, the use of slurs by incompetent speakers, and the contested normative status of 'slurs' as a category. It also suggests a way of dealing with the puzzle of reclaimed in-group

uses. Most importantly, the theory proposed here elucidates the peculiar insolvable nexus of linguistic and political issues that slurs encapsulate.

## Compliance with Ethical Standards

**Conflict of Interest** The author declares that there is no conflict of interest.

## References

Anderson, L., & Lepore, E. (2013a). Slurring words. *Noûs*, *47*(1), 25–48.

Anderson, L., & Lepore, E. (2013b). What did you call me? Slurs as prohibited words. *Analytic Philosophy*, *54*(3), 350–363.

Asim, J. (2007). *The N-word: who can say it, who shouldn't, and why*. Boston, New York: Houghton Mifflin.

Bianchi, C. (2014). Slurs and appropriation: an echoic account. *Journal of Pragmatics*, *66*, 35–44.

Blakemore, D. (2015). Slurs and expletives: a case against a general account of expressive meaning. *Language Sciences*, *52*, 22–35.

Bolinger, R.J. (2017). The pragmatics of slurs. *Nous*, *51*(3), 439–462.

Brandom, R. (1983). Asserting. *Noûs*, *17*(4), 637–650.

Brandom, R. (1994). *Making it explicit: reasoning, representing, and discursive commitment*. Cambridge: Harvard University Press.

Brontsema, R. (2004). A queer revolution: reconceptualizing the debate over linguistic reclamation. *Colorado Research in Linguistics*, *17*(1), 1–17.

Camp, E. (2013). Slurring perspectives. *Analytic Philosophy*, *54*(3), 330–349.

Camp, E. (2018). A dual act analysis of slurs. In Sosa, D. (Ed.) *Bad words: philosophical perspectives on slurs*: Oxford University Press.

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

Cepollaro, B. (2017). When evaluation changes – an echoic account of appropriation and variability. *Journal of Pragmatics*, *117*, 29–40.

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

Cepollaro, B., & Thommen, T. (2019). What's wrong with truth-conditional accounts of slurs. Linguistics and Philosophy.

Croom, A. (2011). Slurs. *Language Sciences*, *33*(3), 343–358.

Farkas, D.F., & Bruce, K.B. (2010). On reacting to assertions and polar questions. *Journal of Semantics*, *27*(1), 81–118.

Gutzmann, D. (2015). *Use-conditional meaning: studies in multidimensional semantics Oxford studies in semantics and pragmatics*. Oxford: Oxford University Press.

Harris, J. (2014). Signaling non-speaker commitment in transparent free relatives: a paired speaker-hearer judgment study. In Rieser, V., & Muller, P. (Eds.) *The Proceedings of the 18th Workshop on the Semantics and Pragmatics of Dialogue* (pp. 71–79).

Hedger, J.A. (2012). The semantics of racial slurs: using Kaplan's framework to provide a theory of the meaning of derogatory epithets. *Linguistic and Philosophical Investigations*, *11*, 74–843.

Hess, L. (2018). Perspectival expressives. *Journal of Pragmatics*, *129*, 13–33.

Hess, L. (2019). Expressive meanings and expressive commitments: a case of meaning as use. In Stalmaszczyk, P. (Ed.) *Philosophical insights into pragmatics*. Berlin, Boston: De Gruyter.

Hess, L. (2020). Practices of slur use. *Grazer Philosophische Studien*, *97*, 86–105.

Hill, J. (2008). *The everyday language of white racism*. London: Wiley Blackwell.

Hom, C. (2008). The semantics of racial epithets. *Journal of Philosophy*, *105*(8), 416–440.

Hom, C. (2012). A puzzle about pejoratives. *Philosophical Studies*, *159*(3), 383–405.

Hom, C., & May, R. (2013). Moral and semantic innocence. *Analytic Philosophy*, *54*(3), 293–313.

Jeshion, R. (2013a). Expressivism and the offensiveness of slurs. *Philosophical Perspectives*, *27*(1), 231–259.

Jeshion, R. (2013b). Slurs and stereotypes. *Analytic Philosophy*, *54*(3), 314–329.

Kennedy, R. (2003). *Nigger: the strange career of a troublesome word*. New York: Vintage.

Lewis, D. (1979). Scorekeeping in a language game. *Journal of Philosophical Logic*, *8*(3), 339–359.

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

Mercier, A. (1995). A perverse case of the contingent a priori: On the logic of emasculating language (a reply to Dawkins and Dummett). *Philosophical Topics*, *23*(2), 221–159.

Morency, P., Oswald, S., de Saussure, L. (2008). Explicitness, implicitness and commitment attribution: a cognitive pragmatic approach. *Belgian Journal of Linguistics*, *22*(1), 197–219.

Nunberg, G. (2018). The social life of slurs. In Fogal, D., Harris, D., Moss, M. (Eds.) *New work on speech acts*. Oxford: Oxford University Press.

Potts, C. (2007). The expressive dimension. *Theoretical Linguistics*, *33*(2), 165–198.

Predelli, S. (2013). *Meaning without truth*. Oxford: Oxford University Press.

Richard, M. (2008). Epithets and attitudes. In *When truth gives out*. Oxford: Oxford University Press.

Ritchie, K. (2017). Social identity, indexicality, and the appropriation of slurs. *Croatian Journal of Philosophy*, *17*(2), 155–180.

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.

Vallée, R. (2014). Slurring and common knowledge of ordinary language. *Journal of Pragmatics*, *61*, 78–90.

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

Williamson, T. (2009). Reference, inference, and the semantics of pejoratives. In Almog, J., & Leonardi, P. (Eds.) *The philosophy of David Kaplan*: Oxford University Press.