dialectica

Vol 75 : 3
September 2021

Neopragmatist Inferentialism and the Meaning of Derogatory Terms —A Defence

Deborah Raika Mühlebach

doi:10.48106/dial.v75.i3.05

Deborah Raika Mühlebach. 2021. "Neopragmatist Inferentialism and the Meaning of Derogatory Terms—A Defence." *Dialectica* 75(3): 1–1. doi:10.48106/dial.v75.i3.05.

Neopragmatist Inferentialism and the Meaning of Derogatory Terms —A Defence

Deborah Raika Mühlebach

Inferentialism seems to be an unpopular theory where derogatory terms are concerned. Contrary to most theorists in the debate on the meaning of derogatory terms, I think that inferentialism constitutes a promising theory to account for a broad range of aspects of derogatory language use. In order to make good on that promise, however, inferentialism must overcome four main objections that are usually raised against Michael Dummett's and Robert Brandom's inferentialist explanations of derogatory terms. This paper aims at debunking these objections and thereby further developing the inferentialist interpretation of derogatory terms. I shall first discuss and reject three of the objections by pointing to the core assumptions of Brandomian inferentialism. Overcoming the fourth objection requires adjusting Dummett's and Brandom's explanation of the meaning of derogatory terms. In order to do so, I shall elaborate on the role that the explication of implicit material inferences plays with regard to different kinds of derogatory terms. The inferentialist account I am proposing fares better in terms of its explanatory power and broadness of application than Dummett's and Brandom's sketchy and oft-criticised views.

1 The Dismissal of Inferentialism for Derogatory Terms

Within philosophy of language, slurs have recently entered the limelight. Slurs are commonly seen as a highly offensive form of derogatory terms. They derogate their targets by virtue of some perceived membership in a social category pertaining to nationality, gender, sexual orientation, race, ethnicity, or religion which is valued negatively. Slurs are both pejorative and derogatory terms. The more general notion of pejorative terms additionally includes words that target people based on individual characteristics or single actions,

for example, "jerk" and "arsehole". These terms are pejorative in the sense that in our discursive practices, we treat them as offensive. The notion of derogatory terms as I use it, refers to terms that structurally derogate their targets, i.e. they are part of and contribute to oppressive social structures.[1] In this paper I aim to focus on structurally derogating terms in the first place. I shall also talk about the offensiveness of slurs and pejoratives. This broad scope allows me to look into mechanisms of change whenever implicitly derogatory terms turn into slurs.

Inferentialism is unpopular where derogatory terms are concerned. Few advocate it to explain the derogatory force of such terms, the variation of force among them, their potential change of meaning over time, or why their derogatory force can be independent of the speaker's intentions. Inferentialist accounts of these phenomena tend to be quickly dismissed, and consequently, there has not yet been much discussion about the extent to which inferentialism is capable of explaining other important aspects of derogatory language use. I think that the rejection of inferentialism is too quick; better than any other theory of derogatory terms, it explains the direct connection between the meaning of our terms and the broader practices in which they are used without losing sight of the semantic differences among various similar terms, such as "sl·t"[2] and "wh·re".[3] An inferentialist account of derogatory terms needs to prove its worth on two fronts: on the one hand, of course, it needs to live up to the constructive task of accounting for various features of such expressions; but it also needs to be able to withstand pressure from a variety of objections that have been levelled against inferentialism—objections that

1 For a detailed discussion of this distinction see Mühlebach (2023a).

2 In what follows, I introduce "·" whenever I use a derogatory term or concept as an example which I do not consider to be part of my vocabulary. Even though certain uses and certain mentions of derogatory terms function non-derogatorily if the context of the utterance makes it sufficiently clear that the speaker does not utter the term in an endorsing way, the mere appearance of derogatory terms may trigger memories of violence. Thus, not semantic but broader political reasons lead me to not spell out the derogatory terms whose content I do not endorse.

3 For example, like perspectivalism (Camp 2013), it accounts for the systematicity of whole perspectives in terms of practical commitments. However, Camp only takes into account whole perspectives without allowing for commitments to specific assertions. Inferentialism, by contrast, preserves the difference between terms such as "sl·t" and "wh·re" which involve the same perspective, but slightly differ in meaning. Whilst both signal the allegiance to a sexist perspective, or even more specifically, the commitment to the claim that the target enjoys sex with too many people, only the latter additionally includes the commitment to a claim about the venality of the target. Explaining the meaning of these two terms only by pointing to a sexist perspective misses the difference in meaning.

have so far proven influential in putting people off inferentialist accounts of slurs.

It is this second, defensive task that I want to tackle in this paper. I leave the constructive task to another paper (Mühlebach 2023b), where I show how inferentialism can account for various distinctive features of derogatory terms, including notably their derogatory (as distinct from their offensive) force which is relatively autonomous from the speaker's intentions and the target's actual feeling of being hurt, the variation of this force among different derogatory terms, the semantic dimension of appropriated uses of derogatory terms by the target group, the fact that in some specific occurrences, the derogatory force of a term scopes out of its embedding, as well as that there are non-appropriated, non-derogatory uses of derogatory terms. Among other things, I argue there that structural derogation involved in derogatory language use is politically more pressing than its offensive potential, that the inferentialist framework does not impose one specific linguistic mechanism onto every type of derogatory expression but allows for various mechanisms to serve as the explanation of different phenomena of derogatory language use,[4] that appropriation by the target group involves semantic change, and that non-appropriated, non-derogatory terms, as well as the scoping out of embeddings, is best explained by making use of the inferentialist distinction between practical endorsement/non-endorsement of inferentially structured commitments, rather the mention/use distinction. But such a positive account, however elaborate, is only as robust as the inferentialist foundation it is built on. One also needs to do the negative work of showing that this foundation does not crumble under the weight of the major objections to it. It is therefore to such a defence of inferentialism that I now turn.

The aim of this paper is to deflect the four main objections raised against an inferentialist theory of derogatory terms. This involves adjusting and further developing Dummett's and Brandom's explanation of the meaning of derogatory terms so that it applies to a broad range of such terms. Accounting for the meaning of derogatory terms in this new way has three advantages: firstly, it enables us to understand using explicitly derogatory and implicitly derogatory terms as two distinct practices, and it explains the workings of

4 For example, it highlights the importance of stereotypical ascriptions as a whole set of (practical) commitments if an expression makes use of a stereotype (such as in "French shower"). Within a broad inferentialist framework we can also make sense of changes in meaning and force over time and explain how formerly pragmatic mechanisms such as derogation through metaphorical force may turn the semantics of a derogatory expression.

doi: 10.48106/dial.v75.i3.05

each. This explanation amounts to situating different kinds of terms on a continuum between explicitly derogatory and purely descriptive terms, which, as I will argue, is crucial to understanding how and why the meaning and force of derogatory terms may change over time. Secondly, and contrary to most theories of derogatory terms, including Dummett's and Brandom's, it provides a general explanation of all kinds of derogatory terms, including gendered, racial, religious, ableist, homo- or transphobic derogatory terms, as well as terms for individual behaviour ("arsehole" or "jerk"). And thirdly, it suggests why criticising derogatory language use by merely making explicit the pernicious inferences that it licenses, often misses the point of the practice of using explicitly derogatory terms.[5]

My negative account complements Lynne Tirrell's (1999) inferentialist view of broader communicative situations in that it shows how such a framework affects how we need to model the semantics of derogatory terms more narrowly. Moreover, it differs significantly from Daniel Whiting's (2008, 2013) view which combines inferentialism with conjunctive non-cognitivism, since my paper clears the path for a full-blown account of the meaning of derogatory terms that does not resort to an additional theory in order to account for derogation through language use. Finally, Esa Díaz-León (2020) aims to defend a semantic strategy by presenting an inferentialist version of Christopher Hom's semantic externalism (Hom 2008, 2012). I take her view to fare better with regard to several objections that have been raised against truth-conditional content views (e.g. Cepollaro and Thommen 2019). However, since her view imports inferentialist ideas into the truth-conditional paradigm, I do not consider it to be a neopragmatist inferentialist account of derogatory terms as I defend it here.[6]

5 The same holds for merely challenging or blocking conventional or conversational implicatures and presuppositions in a specific speech situation.

6 Inferentialism turns the truth-conditional programme upside down by not starting off from reference, but rather from inferentially structured social practices and ultimately explaining reference through inference. Current truth-conditional accounts generally base their theories on the assumption of there being a neutral counterpart to every slur. According to these views, the counterpart and the slur share the same reference. As I have argued elsewhere (Mühlebach 2021), this leads them to ignore a broad range of derogatory terms that lack such a counterpart. Moreover, where *prima facie* applicable, these counterparts are themselves often so complex that the assumption of them being neutral is not true at best and morally highly problematic at worst.

2 Common Objections to the Inferentialist View

Inferentialism receives its name from the assumption that the content of a concept is determined by its inferential relations to other concepts. Concepts are expressed by terms, but a specific term may be ambiguous so that it expresses more than one concept (e.g. "light") and a specific concept may be expressed by different terms (e.g. "red", "*rot*", "*rouge*"). Thus, two tokens of the same term differ in meaning if they express a different concept.[7] In Brandom's pragmatist understanding of semantics, the picture of the inferentialist theory of meaning looks roughly as follows: the meaning of a sentence is determined by the role this sentence plays in the practice of making assertions and giving and asking for reasons. Furthermore, the meaning of a word or an expression is determined by the roles it can play in the assertions of this practice. For a broader understanding of linguistic meaning, we must of course consider a variety of pragmatic moves we can make in and with our speech acts, but as far as semantics are concerned, inferentialism confines itself to the practice of making assertions.[8]

Making assertions and giving and asking for reasons is a social practice that is modelled as a deontic scorekeeping game with different parties. It involves ascribing commitments to other parties as well as undertaking commitments oneself. The most important rule governing such a practice or game is that by making an assertion, you as the speaker undertake the responsibility to provide reasons for your utterance in case you are challenged by a hearer. For the whole practice to work, the different parties involved in this game keep score of both their own commitments and the commitments of the other speakers.

The different sentential roles that an assertion[9] can play depend on the inferential relations of this sentence to other assertions according to the con-

7 In what follows, I shall often use the expression "conceptual norms of discursive communities" in the sense in which conceptual and discursive norms are strongly connected with each other. Linguistic norms govern the use of terms which, in turn, express concepts that are expressed by these terms and are governed by the conceptual norms of the discursive community.

8 Jennifer Hornsby's (2001, 138) objection that the inferentialist model cannot capture how individual speakers use the derogatory term in a particular occasion is misguided. Inferentialists do not hold that particular instances of (derogatory) language use can be explained on the grounds of semantics alone. But if we want to understand why certain words are so apt to be used as weapons while others are not, it is helpful to have a look at their semantics, too.

9 For Brandom, commitments come in the form of assertions, but there are other accounts of inferentialism such as Daniel Whiting's (2007, 2013), which takes non-propositional attitudes to play a similar role in the game of giving and asking for reasons, at least in some cases.

doi: 10.48106/dial.v75.i3.05

ceptual norms of the discursive community (for the following, see Brandom 1994, 157ff.).[10] These relations come in two different forms—in the form of *commitment-preserving* inferential relations, and *entitlement-preserving* inferential relations. As far as the commitment-preserving inferential relations are concerned, as an utterer of an assertion I am committed to further assertions, the so-called *concomitant commitments*. Take the following assertions:

(1) New York is to the East of Paris.
(2) Paris is to the West of New York.
(3) New York is not to the East of Paris.

If I claim (1), I am concomitantly committed to assertion (2), because according to the conceptual rules of the English-speaking community, these two assertions stand in a commitment-preserving relation to each other.[11]

Entitlement-preserving inferential relations, by contrast, do not compel commitments. Whenever I am entitled to make an assertion, i.e. my assertion is taken to be true by other scorekeepers, I am thereby also entitled to commit myself to the claims that materially follow from my assertion.[12] My entitlement to the claim "This is a dry, well-made match" entitles me to be committed to the further claim "It will light if struck." I do not have to be so committed, however, because my first assertion is compatible with the claim "The match is at such a low temperature that friction will not succeed in igniting it" (see Brandom 1994, 169). The interplay of commitment-preserving and entitlement-preserving relations can be put in terms of material incompatibility: being committed to assertion (3) from above, which is incompatible with (1), precludes me from being entitled to claim (2).

10 For reasons of simplicity, I only talk about *the* discursive community and its practical and conceptual norms here. However, our societies consist of several discursive (sub-)communities with somewhat different social practices and, hence, different practical and conceptual norms. Elsewhere (Mühlebach 2021), I model them as communities of practice (cf. McConnell-Ginet 2011; Anderson 2018) and, by the example of the race term "black", I argue that some of the political disputes among different communities of practice amount to semantic contestations of specific terms.

11 Lynne Tirrell (1999, 146ff.) distinguishes between three kinds of commitments: assertional, referential, and expressive commitments. Although the referential and the expressive commitments are semantically reducible to the basic assertional commitments, distinguishing them becomes relevant as soon as we establish a broader inferentialist theory of communication which involves the possibilities of criticising or challenging language use.

12 Brandom's inferentialism concerns material rules (material inferences, material incompatibilities), and not just formal rules. These material rules depend on the content of the concepts involved.

The description of the inferential relations I have given so far highlights the intrapersonal or, according to Brandom, the inter*content* dimension of these relations because the change of normative statuses (commitment/entitlement) produced by an assertion has consequences for the persons already committed or entitled to the claim in question. If I make an assertion, I present it as a reason for other assertions. By putting it forward as a premise from which other assertions can be materially inferred, I change the landscape of further claims I am committed to. Similarly, an assertion that is uttered changes the landscape of additional claims that everyone who is entitled to the assertion in question is now also entitled to.

However, the inferential relations are inter*personally* significant, too. This is where entitlement-preserving relations become especially relevant. If I assert claim (1), I am thereby committed to sentences such as (2) which, according to the conceptual norms of my discursive community, stand in a commitment-preserving relation to (1). In undertaking this commitment, I license or entitle others to attribute that commitment to me. I license them to assume that I endorse both my assertion and the further claims that can be inferred from it. Moreover, I entitle my listeners and others to reassert my assertion, as well as the claims that follow from it.

Inferentialism is often taken to model linguistic exchanges as a rational, co-operative enterprise that is confined to propositional content. This view, however, misses a major part of the inferentialist theory. Inferentially structured propositions are an important part of the inferential web of commitments, but so are non-verbal perceptions and actions. The latter are a part of this web as so-called *language-entry* and *language-exit* transitions (see Brandom 1994, 233–234). According to the first, verbal claims cannot only be made in response to other verbal claims but also in response to non-linguistic, perceptible circumstances. With regard to the second, non-linguistically acting in response to some verbal claim is as much a possible move in the game of giving and asking for reasons as responding by making another claim. For example, cordially hugging somebody right after calling her a b·che are two moves that are as incompatible with each other, just as calling someone a b·che and saying what a lovely person they are would be.[13]

In a more general vein, inferentialists think of semantics in pragmatic terms: sentences are meaningful in the practice of making assertions and giving and

13 This does not apply to reclaimed uses by the target group. According to the inferentialist view, the reclaimed term "b·che" would express a concept different from the derogatory "b·che" because it would license significantly different inferences.

doi: 10.48106/dial.v75.i3.05

asking for reasons. Making assertions includes much more than producing meaningful sound. It involves speakers and hearers both perceptively reacting to the world in accordance with the assertions that are made in this practice and acting upon them correspondingly. This whole practice of making assertions, in turn, is only meaningful as a practice in a broader communicative context. Our broader social norms enable and constrain which conceptual norms our discursive community abides by. The social embeddedness of terms becomes especially important in the case of verbal derogation. A whole set of propositional and practical commitments and, in consequence, the whole social practice surrounding the use of a specific term produce the derogatory force of this term.

Derogatory terms are based on, contribute to, and reinforce social structures of oppression. Social oppression occurs if, through everyday habits and practices, certain people are being systematically disadvantaged and treated unjustly (see Young 1990, chap. 2). Oppressive social structures facilitate or even call for derogatory expressions to name the vertical differences produced by these structures. Such terms, in turn, reinforce these structures: because our discursive practices are inferentially structured, the use of such terms licenses further oppressive speech and action if it is not effectively counteracted.[14] Note that negative evaluations do not per se contribute to oppressive social structures. Some of them are legitimate and harmless, and some are still treated as offensive. Terms that involve legitimate negative evaluations that are treated as offensive are pejoratives. In an egalitarian society there could be pejoratives, but no derogatory terms. However, given that social oppression is so pervasive and multidirectional, it is more than unlikely that egalitarian societies, and thus societies without derogatory language, will ever exist.

Inferentialist approaches to the meaning of derogatory terms are neither new nor undisputed.[15] Many contemporary contributions to the discussion on the semantics and pragmatics of derogatory terms take inferentialism into account, but most of them limit themselves to criticising some of its alleged core assumptions. I take the main objections to be the following four:

(i) *Understanding without endorsement:* Listeners are able to understand sentences that contain derogatory terms even if they do not (morally)

14 For a compelling example, see Tirrell's (2012) discussion of language in the Rwandan genocide.

15 Dummett (1973) uses one example of derogatory terms, Brandom (1994, 2000) adopts it and explains its most basic semantic workings, and Tirrell (1999) develops an account of derogatory speech situations that draws on an inferentialist vocabulary but is not itself semantic.

endorse the inferences that have to be drawn from the utterance of these sentences. Inferentialism cannot explain this condition.

(ii) *Overdemandingness:* Inferentialism seems to overlook the fact that bigoted as well as non-bigoted speakers understand a derogatory term before or even without knowing most of the inferences that arguably need to be drawn from its use.

(iii) *Inability to explain variation in force:* There are subtle differences in pejorative force between different derogatory terms. Inferentialism fails to explain this complex variation of pejorative force, because we cannot assume that competent speakers have knowledge of all the properties that inferentialism takes to be semantically relevant.

(iv) *Indeterminate reference:* Inferentialism faces difficulties in explaining the route from inference to reference when it comes to sentences containing derogatory terms.

As I shall show, the first three objections all call for similar explanations. They can be met by pointing to the inferentialist's core assumptions. The fourth objection, however, requires some extra theoretical work. Whilst objection (iii) applies to slurs or pejoratives specifically, objections (i), (ii), and (iv) concern the inferentialist picture in general. By deflecting these objections, I aim to defend inferentialist semantics in general, and I show that this is helpful in order to understand how the derogatoriness of terms may change over time.

3 Meeting the Objections

3.1 *Understanding Without Endorsement*

Timothy Williamson observes that "we find racist and xenophobic abuse offensive because we understand it, not because we fail to do so" (2009, 141). According to him, inferentialists are not able to explain this. According to Williamson, Dummett believes that speakers and listeners understand the concept B·CHE iff they are disposed to draw inferences according to the following introduction and elimination rules:

B·che-Introduction:	*B·che-Elimination:*
x is a German	*x* is a b·che
x is a b·che	*x* is cruel

doi: 10.48106/dial.v75.i3.05

These rules specify the term's inferential role in a specific language by using the vocabulary from this language without the term in question. The introduction rule states under which circumstances it is appropriate to apply the term in question. The elimination rule, in turn, states the consequences of applying the term. Hence the rules show how the term in question can be "introduced" to a language and then "eliminated" from its vocabulary so that we get a picture of the term's inferential role without resorting to the term itself.[16] Since, according to Williamson, non-xenophobic people who are not willing to draw the inferences from the "b·che" rules nevertheless do understand the term "b·che" perfectly well, he doubts that it is the disposition to draw these inferences that makes speakers and listeners understand the term.

This objection rests on two mistakes. First, inferentialists would not agree with Williamson's claim that non-bigoted speakers are not willing to draw the relevant inferences. Being disposed to draw inferences is not a matter of will. Rather, I fully understand a concept iff I know what inferences are licensed by its use, i.e. iff I know which inferences are correct according to the conceptual norms of the discursive community. The crucial question with regard to derogatory terms is whether I think that the inferences which are correct according to the conceptual norms of the discursive community are also morally and epistemically correct. Williamson's objection conflates the two standards of correctness, the semantic and the moral and epistemic standard. We evaluate somebody's understanding of a term with reference to the semantic standard of correctness, whereas the moral and epistemic standard leads people to criticise the use of certain terms and to refrain from using them. Understanding whilst refraining from using a term amounts to both knowing the inferential role of the term in question and rejecting a whole set of practices in which these inferences are treated as correct.

Thus, secondly, Williamson's objection conflates understanding a concept with using this concept. People can smoothly communicate even if they have different moral and epistemic standards because *understanding* a term does not necessarily imply endorsing its concomitant commitments, whereas *using* it necessarily implies endorsement.[17] Listeners may well know the

16 Dummett's prime example is the introduction of the logical connective "&": the introduction rule for "&" is the transition from "p, q" to "p&q" and the elimination rule is the transition from "p&q" to "p, q".

17 This difference can also be stated in terms of *attributing a commitment* to another person vs. *undertaking or adopting the commitment* oneself. See Brandom (2000, 169).

inferences that can be drawn from assertions that involve the term "b·che", whilst refraining from endorsing some or even all of these inferences. If a listener understands but does not endorse the assertion in question, the utterer of the assertion is committed to the inference "*x* is more likely to be cruel than other Europeans" while the listener is not. However, the listener knows that the utterer is so committed. Yet the listener does not think that the utterer is entitled to the inference, since the listener herself does not endorse the inference. Thus, by using certain terms, the speaker is committed to the inferences that must be drawn from the utterance in question and he entitles his listener to draw these inferences, too. The listener, however, needs neither to be committed to these inferences, nor must she think that the speaker is entitled to draw these inferences if she does not endorse them.[18] If the listener does not endorse a certain inference, it is because she thinks that the inference in question is not sound, i.e. the assertion that is supposed to follow is not true and, in the case of derogatory terms, that it is morally and/or politically problematic.

3.2 *Overdemandingness*

Critiques of inferentialism hold that it is far too demanding with regard to the cognitive work speakers and listeners would have to undertake if they were really committed to all the inferences ascribed to them. Take the British boy Sebastian telling his friends in the aftermath of WWI:

(4) There are too many b·ches in town these days.

Some listeners of (4) may easily understand that Sebastian's utterance is derogatory without exactly knowing what claims he is thereby committed to.[19] The problem is best captured by Hornsby who already leads us partially out of it:

18 As will become clear in my response to the objection of indeterminate reference, my version of an inferentialist account views the introduction rules as they are stated here as incorrect. According to the introduction rules that I will propose, my response to the understanding without endorsement objection amounts to the following view: if a listener understands a term, they know under which circumstances the bigoted speaker applies the term even if the listener thinks that the conditions of application do not obtain. And the listener also knows what consequences the use of the term will have even though the listener does not endorse these consequences. With regard to many strongly derogatory terms, this view amounts to the null-extension view as it has been proposed by Hom and May (2013, 2018).

19 See Jeshion (2013, 245) for a similar view.

doi: 10.48106/dial.v75.i3.05

> With enough hard work, a historically minded social theorist might [...] provide the inferentialist with materials to make explicit that which is to be treated as implicit in each racist word. The question now is whether speakers themselves undertake commitments, which the historian uncovers. The answer appears to be *No*. [...] We are trying to account for something readily picked up by speakers of a certain social formation; and we have to allow for the fact that racist and other derogatory words can be passed on quite easily. If speakers' involvement with the ideology went as deep as it would need to in order to be implicit in their very use of words, then common understandings would be difficult to preserve. (Hornsby 2001, 137)

In short, the worry is this: there is a basic understanding of terms prior to many inferences we might draw from their use.

Two things have to be noted here. First, Hornsby confounds the inferentialist's use of "commitment" with an internalist understanding of "speaker meaning". Which claims the utterer of (4) is committed to, is not a question of conscious inferences made by the speaker in the first place, but a question of the term's inferential role in a discursive community. According to the inferentialist view regarding derogatory terms, it is indeed the case that historians, social scientists, and people with a sense of social mechanisms are the ones most likely to uncover the commitments of speakers in a discursive community. Only by making the commitments that are implicit in our discursive practices explicit do we see whether our commitments correspond to what we think we are saying. Sometimes our conceptual commitments are not so much in agreement with what we mean to say.

This, however, might give us an awkward picture of concept use in everyday communication. How should people understand each other if they do not know any of the important inferences to which they are committed? Thus, the second point is that conceptual understanding comes in degrees. If Sebastian's listeners do not take him to be committed to any of the assertions he is actually committed to, we would simply say that they do not understand the concept B·CHE at all. As soon as the listeners have some grasp on the basic inferences, they understand that the utterance is derogatory. In that case, however, they have a very poor understanding of the concept of b·che. They are not able to see that the alleged despicableness is explained by the likeliness to be cruel and not, for example, by some ascribed special visual appearance. A simple

distinction between understanding a concept and not understanding it is untenable. Various people are in different states of conceptual competence with regard to different concepts, but the difference between those states is not a distinction of kinds, but of degrees (Brandom 1994, 120).[20] The more inferential roles of a concept I know, the richer my understanding of the concept in question is. And the richer my understanding of the concept I use is, the more I move away from merely deploying a randomly picked up concept towards using it with the explicit knowledge about my linguistic and extra-linguistic commitments that come along with it.

It is certainly a helpful first move to attend to degrees of understanding in order to account for different understandings of derogatory terms by different people. However, we need not forget that slurs are part of politically relevant language use. Since politics are about collectively negotiating (social) realities, the contents of words that name these realities are often part of these negotiations, too. It is a constitutive part of such language use that people might disagree about correct uses in general or in different contexts. Thus the political complexity of such terms sometimes runs contrary to the interest of philosophers of language in always exactly determining the truth conditions of a sentence, or the exact set of inferences that come along with an assertion.

Moreover, the underlying complex and dynamic social structures of derogatory language use make it likely that many derogatory concepts function as cluster concepts, as has been suggested by Croom (2011, 353ff). With regard to the term "ch·nk", for example, a full understanding of the term does not involve knowing about the long list of all the stereotypical ascriptions for Chinese people, such as being slanty-eyed, devious, good at laundry, etc., but contextual norms make it clear which parts of the whole set are relevant to a specific speech situation. However, stereotypes of a social group often systematically hang together and are part of a whole set of social meanings that affect the meaning of a term in a specific context (see Mühlebach 2022). Language users who are aware of the connections between these social meanings might well be taken to possess a deeper understanding of the terms they use than those who are not.

I have argued elsewhere that inferentialism is in a good position to account for the social complexities of such words (Mühlebach 2021). For example, it allows for different communities to have different uses of a specific term, as

20 Cf. Higginbotham (1998, 149ff.) for a helpful distinction of different states of conceptual competence.

doi: 10.48106/dial.v75.i3.05

Luvell Anderson (2018) has argued regarding the n-word. Moreover, inferentialism provides the resources to theorise about the power relations among these communities and how these affect the meaning of such terms. As I have shown, that is true even if we take many slurring concepts to function as cluster concepts.

3.3 *Inability to Explain Variation in Derogatory Force*

Hom worries that inferentialism cannot explain why "speakers have a pretty good understanding of the lexical, negative ordering of slurs (e.g. the n-word is worse than 'ch·nk' is worse than 'l·mey', etc.)" (2010, 175, "·" added by the author). Since the list of licensed inferences would need to distinguish properties in very subtle ways and we cannot take competent speakers to have knowledge of all these properties, inferentialists are not able to explain the difference in force between, say, the n-word and "ch·nk".

Hom's insights into the variation of force are important, but, as I shall argue here, the variation can be equally accommodated by inferentialist semantics. In Hom's view,

> the derogatory content of an epithet is semantically determined by an external source. The plausible candidates for the relevant external social practices that ground the meanings of racial epithets are social institutions of racism. [...] An institution of racism can be modelled as the composition of two entities: an ideology, and a set of practices. (2008, 430–431)

If we state Hom's claims in inferentialist terms, his explanation of derogatory force fits well into the broad inferentialist picture provided by Brandom or Tirrell. In a pragmatist understanding of inferentialism, the meaning of a term is determined externally through the inferences that are treated as valid by the discursive community.[21] Recall my sketch of the inferentialist framework called *deontic scorekeeping* from above: the content of a term is determined by the term's inferential role in the game of giving and asking for reasons. The parties involved in this game keep score of what the other parties and they themselves are committed and entitled to. Every assertion

21 Note that treating inferences as valid ones is not exhausted by those that are consciously treated as such. Inferentialism leaves room for the fact that people, individually or even collectively, are sometimes not aware of what they are doing. Explicating their implicit inferences amounts to bringing these inferences to the realm of reasoning.

that is made in this game alters the scores of the scorekeepers. In other terms, every assertion changes the set of claims to which the parties involved in the game are inferentially committed and entitled.

Scorekeeping is a social game. Which inferences count as correct depends on the social norms of the discursive community. The stronger the racism against a social group within a discursive community, the more numerous the racist inferences that are licensed by this community, and the more pernicious. The same holds true for sexism, xenophobia, and other discriminatory institutions. While the war-experienced French community directly after WWI would have been likely to treat the inference from "*x* is of German nationality" to "*x* is likely to be cruel" as correct, this is not what French communities would treat as a correct inference today.

In Hom's explanation of the force of derogatory terms, a set of specific social practices and a web of negative ideological beliefs add up to the social institution of racism. Since inferentialism not only includes propositional commitments but extends to language-entry and language-exit transitions, non-verbal perception and action, which are important for Hom's view, are not missing in the inferentialist account. Inferentialism captures the finely-grained ordering of different derogatory terms by the *set of commitments* that come along with these terms, and not by single inferences such as moving from "*x* is a b·che" to "*x* is likely to be cruel." The variation of derogatory force is dependent upon the actual devaluation of the target in the practices in which the expression is used.[22] How strongly a target can be derogated by the content of a concept depends on the power relations within a given discursive community, i.e. on how much a target can be devalued according to their social, economic, and cultural capital within the discursive community.

In cases of strongly derogatory terms such as the contemporary meaning of the n-word, the set of commitments includes claims such as that black people are inferior to white people, as well as practical commitments to treat black people in dehumanising ways. Commitments are not necessarily consciously held beliefs, but rather commitments that are operative in a given practice

22 Note that the offensive force and the derogatory of an expression are not necessarily the same. Offensiveness tells us something about which utterances a discursive community treats as appropriate, the derogatory force varies according to the actual devaluation of the target which is operative in the economic, social, and cultural practices in which the use of the expression is embedded. Elsewhere (Mühlebach 2021), I argue that criticism of language use often consists in that a sub-group of a discursive community tries to make other members of the community aware of the actual derogation involved in the use of a specific expression.

doi: 10.48106/dial.v75.i3.05

and that can be made explicit if we try to rationalise the practice. As we know from critical race theorists such as Linda Alcoff (2015), contemporary race relations in Anglo-American countries are such that whiteness functions as the unmarked norm whereas other races function as categories that deviate from it.[23] Thus, to date, the alleged inferiority of black people eventually amounts to their alleged inferiority to the white race. As Alcoff shows, it is a complicated question as to whether race relations can be transformed in a way such that in a distant future, whiteness could simply be one category among many others.

Given the role that the n-word historically played in practices of slavery, lynching, and other types of inferiorisation and dehumanisation, and given that the power relations between whites and blacks are still highly asymmetrical, the current use of the n-word still carries the weight of inferiorising black people and licensing to treat black people cruelly. By contrast, the US-American use of the term "l·mey", which targets British people, licenses inferences that are much weaker in their derogatory force. In the US, there is no deeply rooted xenophobia against British people and there are no clear hierarchical power relations in play on which a web of deeply pernicious inferences could be based. If it is used at all, it licenses mocking behaviour of the British stereotype.

With regard to two or more expressions that target the same group of people, their derogatory force may still vary. For example, even though the n-word and "d·rkie" are both highly derogatory, the practices in which they have typically been used differ. The former involves dehumanising, inferiorising, and especially violent and cruel behaviour, whereas the latter also involves inferiorising, but rather patronising than cruel behaviour. At least, that is true for their historical use. The more they are used in the same practices and in the same way, the more their derogatory force, as well as their meaning, converges.

3.4 *Indeterminate Reference*

Both Williamson and Hom worry that the inferentialist faces the problem of unfixed references when explaining the meaning of derogatory terms. By

23 Note, however, that this does not imply that whiteness is also unmarked for the ones who are dominated by whites. Sara Ahmed rightly observes that "whiteness is only invisible for those who inhabit it. For those who don't, it is hard not to see whiteness; it even seems everywhere" (2004)

responding to their worry, I shall address their misunderstanding of Brandom and Dummett: in the case of derogatory concepts they assume a symmetrical relationship between introduction and elimination rules of concepts, whereas Brandom and Dummett take this relationship to be asymmetrical. My response in this section will be Brandomian in order to show what we are able to explain with the help of classical inferentialism. In section 4, however, I shall point to the limits of this Brandomian view, and in section 5 I shall show that the distinction between explicitly and implicitly derogatory terms is crucial to an explanation of meaning in derogatory language use.

Recall the explanation of a term's inferential role in a game of giving and asking for reasons. This role is best explained using a vocabulary from the same language that does not contain the term itself. We introduce the term into this language by stating the circumstances under which its application is appropriate (the introduction rule). And we state the consequences of applying the term (the elimination rule). According to Dummett, the "b·che"-introduction rule is the transition from "*x* is German" to "*x* is a b·che" and the elimination rule is the transition from "*x* is a b·che" to "*x* is cruel." Williamson's and Hom's objection rests on the principle according to which expression *E* refers to *X* if "the hypothetical assignment of *X* as the reference of *E* makes R(*E*), [i.e. the rules for the use of *E*] truth-preserving" (Williamson 2009, 143). Williamson and Hom convincingly argue that there "is no determinate route from inference to reference" (Hom 2010, 175) because, according to the introduction and elimination rules of "b·che", there is more than one set of objects to which "b·che" can refer.

Instead of spelling out in what ways the reference might be indeterminate, let me note that the objection rests on the assumption that the relationship between the introduction and the elimination rule of derogatory terms is symmetrical. However, both Dummett and Brandom treat derogatory sentences as clear examples of an asymmetry between those two rules. The difference between a symmetrical and an asymmetrical relation between introduction and elimination rules lies in that only with regard to the latter, the elimination rule introduces new inferences which were not yet part of the web of inferences that are relevant for the circumstances of application. Terms that are stably institutionalised and not just about to change do generally have introduction and elimination rules which are *in harmony* with each other, i.e. they are *symmetrical*. If we introduce a new term into an existing language and if its elimination rules are part of the web of inferences which already determines the introduction rules, this is called a *conservative extension* (see

doi: 10.48106/dial.v75.i3.05

Williamson 2009, 138–139). A *non-conservative extension*, by contrast, is built on the *asymmetrical* relation between these rules. In this case, the elimination rule involves further inferences which have not been part of the web of inferences relevant for the introduction rule yet. As we will see, conceptual changes in everyday language use and conceptual progress in science are phenomena in which we are confronted with asymmetry between the introduction and elimination rules.

The wide-spread confusion about symmetrical and asymmetrical introduction and elimination rules in the case of derogatory terms makes the indeterminate reference objection *prima facie* pointless because it targets a view that neither Dummett nor Brandom adhere to. However, there is something peculiar about the fact that all critics of the inferentialist position who are concerned with the reference problem take the introduction and elimination rules in the case of "b·che" to be symmetrical, even though Dummett and Brandom clearly argue for their asymmetrical relationship. In the remaining part of this paper, I shall explore in what ways this confusion can be made fruitful to adjust Dummett's and Brandom's insufficient account of a broad range of derogatory terms. Dummett and Brandom assign different roles to the asymmetry of introduction and elimination rules with regard to "b·che". Take Dummett's remarks on these rules first:

> It [the distinction between introduction and elimination rules] remains, nevertheless, a distinction of great importance, which is crucial to many forms of linguistic change, of the kind we should characterize as involving the rejection or revision of concepts. Such change is motivated by the desire to attain or preserve a harmony between the two aspects of an expression's meaning. A simple case would be that of a pejorative term, e.g. 'B·che'. (1973, 454, "·" added by the author)

Claiming an asymmetrical relation of introduction and elimination rules with regard to 'b·che' is to say that "*x* is likely to be cruel" stands in no inferential relation to "*x* is German," or put differently, "*x* is likely to be cruel" is not part of the inferential web of assertions which is relevant for utterances of the form "*x* is German." In order to attain harmony between the introduction and elimination rules, one of them, or both, have to be revised or the concept needs to be rejected. Dummett thinks of B·CHE as an example of concepts which need to be rejected for reasons of harmony: there is no way in which

non-xenophobes would and should be licensed in drawing the inference from "*x* is German" to "*x* is disposed to be cruel."[24]

Brandom reminds us, however, that this lack of conceptual harmony can be productive and does not lead to the rejection of concepts *per se*. According to him, the lack of harmony between the introduction and elimination of B·CHE is analogous to the lack of harmony in cases of conceptual change. A prime example of conceptual change is GRAVITY before and after Einstein's theory of relativity. The Newtonian introduction rules were not in harmony with the elimination rules anymore, as soon as Einstein put forward his theory of relativity. The inferences of the elimination rule were not yet part of the inferential web of the introduction rule. According to Brandom, B·CHE works similarly: "the problem with 'B·che' or 'n·gger' is not that once we explicitly confront the material inferential commitment that gives them their content, it turns out to be *novel*, but that it can then be seen to be indefensible and inappropriate" (1994, 127, "·" added by the author).

Unfortunately, Brandom does not mention that in the case of GRAVITY, the introduction rules *do* change once we accept the new consequences (or elimination rules) that the application or use of this concept has, so that introduction and elimination rules ultimately come into harmony again. The concept of gravity could not be introduced with Newton's criteria anymore as soon as Einstein's theory of relativity, which made explicit the new inferences that had to be drawn from the use of this concept, was widely accepted. The introduction rules of GRAVITY had to be adjusted once the conceptual change brought about by scientific progress had been made explicit.

The same applies to conceptual changes and conceptual clarification in everyday language use. Take the example of RAPE. For a long time, there was no way in which, according to the predominant linguistic culture, rape could have happened between a married couple.[25] It was only with the rising awareness of women having sexual rights independently of their marital status that non-consensual sexual activity towards one's wife could be socially and legally treated with the same consequences as non-marital rape. In order to attain harmony between the introduction and elimination rules of RAPE,

24 Or, as Brandom's (2000, 70) remarks on Oscar Wilde's trials where he was accused of blasphemy by the cross-examining Mr. Carson suggest, Wilde gave *the only answer he could give* by saying " 'Blasphemous' is not a word of mine." For a more detailed description of the events of Wilde's trials, see Montgomery Hyde (1962, 121ff.).

25 Legal and dictionary definitions were such that "rape" applied only to a man's penetrating a woman who was not his wife. See McConnell-Ginet (2006).

doi: 10.48106/dial.v75.i3.05

the former introduction rules had to be revised so that now, it is appropriate to apply the term "rape" whenever sexual activity and usually sexual intercourse without the other person's consent is concerned, regardless of the relationship between the persons involved.

If we take Dummett as seeing the difference between endorsing and rejecting inferences once they are made explicit, and if we take Brandom to agree that in cases such as the conceptual change of GRAVITY, the introduction rules had to be adjusted once the changed material inferences were made explicit, their views on introduction and elimination rules do not differ significantly. It can be formulated as follows: if the set of claims that are the consequence of applying or using a concept is not yet part of the set of claims that license the use of this concept, this is a non-conservative extension. According to this picture, GRAVITY and RAPE on the one hand, and B·CHE on the other are examples of non-conservative extension with different moral consequences. In the first case, making the (new) commitments explicit leads to the revision of these concepts by adjusting their introduction rules. In the case of the latter, making the commitments explicit and, as a consequence, adjusting the introduction rules leads non-bigoted speakers to reject its use altogether. In both cases Williamson's and Hom's worry does not apply anymore. Once the introduction and elimination rules are in harmony again, the reference is determinate.

The idea behind making explicit the material inferences that are implicit in our discursive practices is the following: our language use is governed by the conceptual norms that are embedded in our broader social practices. The inferences that are licensed by our everyday use of concepts are material, not formal inferences. With regard to many of our concepts it is not obvious what the circumstances of their appropriate application and the consequences of their use are. The point of explicating the inferences hitherto implicit in using GRAVITY or RAPE, i.e. the claims to which we are committed according to the conceptual norms of our discursive community, is to bring these implicit inferences into the realm of reasoning. Once these commitments are on the table, they can be criticised or justified, and the concepts can be rejected or revised. In cases of revision, we adjust the introduction or elimination rules so that they are again in harmony with each other, at least as long as further relevant material inferences are made explicit.

There are several historical cases of derogatory language use in which the explication of the formerly implicit inferences has led to a widespread rejection of the concepts in question. Take CH·NAMAN and the US American

use of OR·ENTAL as two examples of such implicitly derogatory concepts. For a long time, their introduction rules were considered to be the moves from "Chinese man" to "Ch·naman" and "person from a Near or Middle Eastern country" to "Or·ental". But their underlying social practices were formed in such a way that most of their uses committed the speakers to devaluating and exoticising inferences.[26] Only with the rising awareness of these underlying social structures was it possible to make explicit the formerly implicit inferences. As many English speakers do not openly approve of the devaluating and exoticising conceptualisation of people from China, Asia, or the Near and Middle Eastern countries, the use of these terms has diminished remarkably and so made room for alternatives such as "Chinese person".[27]

4 Broadness of Application

Brandom holds that his remarks on the concept of b·che "should go over *mutatis mutandis* for pejoratives in current circulation" (2001, 86). Moreover, he thinks that what makes this French epithet from WWI especially suitable for semantic investigations is that "we are sufficiently removed from its practical effect to be able to get a theoretical grip on how it works" (2001, 89). I doubt that "b·che" does the work Brandom expects it to do. The explanation does not generally hold true for all kinds of derogatory or pejorative terms such as "sl·t" or "arsehole", nor does this explanation help us understand the phenomenon of explicitly derogatory terms, including "b·che" itself. Our temporal and linguistic-cultural detachment from its usage in WWI, I contend, is exactly the reason why it is difficult to understand the basic workings of this term's use.

A closer look at the practice of expressing explicitly derogatory concepts suggests that not every derogatory concept is defective in the way Brandom assumes and thus it is questionable whether we know enough about the phenomenon of derogatory or pejorative terms in general if we understand the way in which the concept B·CHE is defective. Our English vocabulary makes

26 See Stavroula Glezakos (2013) for an illuminating discussion of "Ch·naman".

27 Contemporary examples of terms that are widely seen as purely descriptive, categorising social terms are "woman" or "black person". As feminist and critical race scholars show, however, their use is still governed by more or less hidden sexist and racist practices. This does not mean that they have to be rejected altogether, but rather that both a reconceptualisation and a change of the underlying social structures are necessary—talking about persons-gendered-as-woman and persons-racialised-as-Black is an example of the former; feminist and anti-racist political contestations are examples of the latter.

doi: 10.48106/dial.v75.i3.05

use of various kinds of derogatory terms targeting gender, race, ethnicity, nationality, religion, sexual orientation, ability, physical appearance, single actions, or patterns of behaviour. As sexism, racism, xenophobia, ableism, homo- and transphobia are institutionalised differently, they come in different forms and in different degrees of derogatoriness. In light of their variety, the conclusion that derogatory or pejorative terms always express defective concepts is too hasty.

Take, for example, the term "arsehole". It is undisputed that this term is a pejorative. Nevertheless, I am sceptical of both its derogatoriness and defectiveness. If we take an utterer of "*x* is an arsehole" to be committed to claims along the lines of "*x* arrogantly allows himself to enjoy special advantages," there is nothing defective to be found here. For if we follow Aaron James (2012, 205ff.) in his analysis of "arsehole", we see that the use of the term is mostly gendered—hence the "allows *him*self" from above—in that our unjust social arrangements make it possible mostly for men to arrogantly allow themselves to enjoy special advantages. Thus, even though "arsehole" is pejorative, i.e. we treat it as an offensive term, it is not derogatory. It is not based on nor contributes to oppressive social structures. On the contrary, it might even be a suitable means to point to behaviour that takes advantage of unjust social structures. Moreover, although we might always be able to put our message in a less forceful and more diplomatic way than by using the pejorative term "arsehole", it does not express a defective concept. All of the inferences involved are valid.

However, many of the pejorative and derogatory concepts that are currently in circulation are indeed defective, and thus Brandom's analysis of B·CHE might still prove valuable to understanding those. But even among the concepts that license flawed inferences, we find a considerable number that do not fit the model of "descriptive" circumstances of application and "evaluative" consequences of use proposed by Brandom's discussion of the B·CHE example.[28] Take the case of gendered derogatory terms such as "sl·t", "b·tch", or "S·ssy". Lauren Ashwell (2016) has argued that gendered slurs do not have "purely descriptive" and unproblematic circumstances of application. The circumstances of application of "sl·t", for example, already contain the evaluative description "*x* has sex with too many partners." If we follow Ashwell,

28 Brandom notes that "[a]lthough they are perhaps among the most dangerous, highly charged words—words that couple 'descriptive' circumstances of application with 'evaluative' consequences of application—they are not alone in incorporating inferences we may need to criticize" (2001, 87).

our everyday language does not seem to provide any non-evaluative way of picking out who the target of the assertion "*x* is a sl·t" is.[29] Thus, Brandom's explanation of the "b·che" case does not seem to apply to gendered derogatory terms, either.

The non-applicability of Brandom's explanation of "b·che" to other common derogatory terms such as "arsehole" and "sl·t" calls for a re-examination of his account of derogatory terms. Its fruitfulness for understanding cases such as "Ch·naman" and "Or·ental" notwithstanding, it does not fully apply to the phenomenon of explicitly derogatory terms such as the n-word, "sl·t," "ch·nk", "k·ke", or "arsehole". Moreover, it even does not rightly capture the case of "b·che" because it conflates two distinct practices—the use of explicitly derogatory and the use of implicitly derogatory terms.

5 Implicitly vs. Explicitly Derogatory Terms

As discussed above, terms such as our historical examples "Ch·naman" and "Or·ental" have not always been considered pejorative. They were institutionalised within the English-speaking discursive community, which has been dominated by white people, in a way that has treated them as purely descriptive terms. However, the dominant use of these terms functioned derogatorily, and with the rising awareness of the devaluating and exoticising conceptualisation of Ch·namen and Or·entals, the formerly implicit inferences could be made explicit. By bringing them into the realm of reasoning, people had to take a stance by either openly committing themselves to pernicious inferences or by refraining from using the terms and engaging in practices of using alternative terms. The terms that philosophers of language are usually concerned with, by contrast, are explicitly derogatory terms such as the n-word, "sl·t," "ch·nk", or "k·ke". To say that these are explicitly derogatory terms is to say that their use is put under strict social constraints and that more or less competent English speakers know that by using them, they are strongly derogating their target. The way in which these terms are used among bigots, or to hurt somebody, and the way their use is sanctioned suggest that users of such terms know both that and in what sense these terms are derogatory.

If people roughly know about the *that* and *how* of the derogation involved, the main inferences that have to be drawn from the use of derogatory terms

29 Note that if we find a non-evaluative description, such as in a sociological or meta-language vocabulary, for some gendered slurs, then the objection of indeterminate reference, as it has been raised in the case of "b·che", could be raised against the case of the respective gendered slur.

doi: 10.48106/dial.v75.i3.05

are explicitly available in the realm of reasoning. Whilst in the case of other politically significant terms such as our historical example "Ch·naman" and contemporary uses of "woman", "disabled", or "immigrant", we still need to work out in what relation to derogatoriness these stand, the use of the n-word, for example, is institutionalised in a way in which people who use it both know that it is not used synonymously with "black person", and that to call people that term is to ascribe racial inferiority to them and not, for example, the predisposition to cruelty. Similarly, competent French speakers in the period surrounding WWI knew that with the use of "b·che" they were derogating their target because of the (alleged) predisposition to cruelty and not because of a physical appearance held to be different from a specific norm. And competent English speakers know that with the term "sl·t", they are derogating their target because of (a behaviour that indicates) having sex with too many partners and not because of being too assertive.

But if we take the main inferences that are licensed by the use of explicitly derogatory terms to already be more or less explicit, this affects the introduction rules of the terms in question. There is no reason as to why the cases of the n-word, "ch·nk", or "b·che" should be different from making explicit the inferences with regard to the concepts of gravity, temperature, and rape. Recall the case in which the changed consequences of the use of GRAVITY have been made explicit once Einstein's theory of relativity was fully established. By accepting these new consequences of use, the circumstances of application had to be adjusted, too. To continue with introducing the concept of gravity with Newton's old criteria, which are not in harmony with Einstein's consequences of use, is pointless if there are new criteria available according to which the circumstances of application that are in harmony with Einstein's consequences of use can be framed.

Applied to the explicitly derogatory n-word, the circumstances in which for a speaker, the application of this term is appropriate are those in which the speaker is already committed to the claims "*x* is a black person" and "*x* is inferior to white people" (among others). Similarly, the circumstances of application in the case of "b·che" would not only involve "*x* is German," but also "*x* is likely to be cruel." These "new" commitments do not merely add the speaker's evaluative attitude towards the target to the descriptive content of the term, rather they shape the content of what the speaker is talking about. This interpretation brings the workings of racist and xenophobic terms closer to those of sexist terms ("sl·t") and terms for individual behaviour

("arsehole") in that all of them require evaluative descriptions to be part of their circumstances of appropriate application.

If we introduce evaluative descriptions into the circumstances of application, this has at least two benefits, besides enabling a better understanding of the bigoted practice of using derogatory terms. Firstly, we can make sense of the inferentialist claim that assertions function as premises and consequences of inferential moves in the game of giving and asking for reasons. Scorekeepers draw the inferences that are licensed by the speaker's assertion. Thus, the assertion serves as a premise for further moves in the scorekeeping game. But it is also a consequence of former claims made in the game of giving and asking for reasons. By requiring the conditions of application of, say, the n-word to include "*x* is black" and "*x* is inferior to white people," we can explain why the assertion "*x* is a n·gger" leads to puzzled reactions on the part of a listener both in the case in which the speaker is talking about a Mexican person and in the case in which her record of assertions commits her to claims such as that race is socially constructed and that people are to be treated equally regardless of their race.[30] In both cases the circumstances of application are not given so that making sense of the speaker's assertion becomes difficult.

Note that we can judge whether the circumstances of application are given according to two different standards: the semantic and the epistemic and moral standard. Given my own set of commitments, a bigoted speaker may correctly apply the n-word in a specific situation in the sense in which they correctly use the term according to the conceptual norms of N·GGER. They meet the semantic standard. However, since I take most of their inferentially organised web of commitments to be both epistemically and morally false, there is, according to my view, no situation in which the circumstances of application of the n-word are given. In this regard, the inferentialist position is the same as Hom and May's (2018) null-extension view: most derogatory terms do not refer to anything in our world because they involve commitments to epistemically false claims.

The two standards allow us to distinguish between different types of language critique. On the one hand, we can criticise uses of terms if they violate the conceptual norms of the concepts they aim to express. Just as in my example from above, we can point to semantically false uses of the n-word because

30 Just as in the incompatibility case of using "b·che" and hugging the target before, this does not apply to the reclaimed use of the n-word by some members of the target group.

doi: 10.48106/dial.v75.i3.05

we know that the speaker must not be committed to the claims "*x* is Mexican" or "*x* is racialised-as-Black" if they want to apply the n-word according to the norms of the bigoted discursive community. These claims are not part of the elimination rule determining the content of the n-word. On the other hand, we can criticise the use of the n-word not only in specific situations, but in general. Its use always commits us to epistemically false and morally problematic claims, so that we should refrain from using it altogether if we want to make epistemically correct and morally unproblematic claims about the world.

The second benefit of including evaluative descriptions into the circumstances of application is that the indeterminate reference objection that has been raised against the inferentialist position becomes obsolete. Williamson and Hom have worried that the term "b·che" does not definitely refer if we take its introduction rule to be the transition from "*x* is German" to "*x* is a b·che" and its elimination rule the transition from "*x* is a b·che" to "*x* is likely to be cruel." By treating explicitly derogatory terms as expressing concepts whose main inferences are already made explicit in the realm of reasoning—which is why they are treated as derogatory terms—we must change their introduction rules so that they are in accordance with their elimination rules again, at least as long as the meaning of these terms is not about to change, or as long as we do not learn about new implicitly operative inferences. Above I have shown what this looks like regarding the examples of "b·che" and the n-word. If the leap between the circumstances of application and the consequence of use is bridged this way, the elimination rule of a specific derogatory term does not add any new claim that is not yet part of the web of inferences relevant to the introduction rule. Hence, the objection of unfixed reference does not apply anymore.

If we acknowledge that the use of explicitly derogatory terms is stably institutionalised in our discursive practices, we need to explain not only why we should refrain from using many of them, but also how their usage does not create communicational impasses. Many people still use the n-word and as wrong as they are in doing so, they do it quite successfully. Most importantly, their use of this term does not necessarily lead them to materially incompatible commitments. Bigoted speakers do not use the n-word *despite* the fact that it commits them to the claim that black people are inferior to whites. They use it *because* they are so committed. If we seek to understand how the n-word, "ch·nk", or "sl·t" are correctly used, we need to turn to the commitments of their users, and not to the commitments of people who refrain from using

these terms, because it is the former who uphold the linguistic practice of consistently using such derogatory terms.

The fact that many derogatory terms are consistently used, however, suggests that making explicit pernicious inferences is but a start in the enterprise of criticising derogatory language use. Merely making pernicious inferences explicit in the cases of "n·gger", "f·ggot", or "k·ke" does not bring us far if our opponent disagrees about them being pernicious. In this case, the criticism has to be more sophisticated and amounts to criticising a whole set of social practices. Other cases such as the use of "arsehole" call for yet another kind of criticism. Making explicit the inferences involved in its use, I contend, does not tell us anything about why we should refrain from using it on certain occasions (or altogether?).

Thus, on the one hand, the explication of implicit inferences is crucial to criticising derogatory language use, but on the other, it does not do all the work that is required for different kinds of derogatory terms. By taking into account a variety of implicitly and explicitly derogatory terms and finding common mechanisms for their basic workings, the version of inferentialism about derogatory terms that I have proposed enables us to understand that explicitly derogatory terms differ from implicitly derogatory terms only in degree. Most theorists who work on the meaning of derogatory terms set out to account for derogatory terms based on their divergence from their allegedly purely descriptive correlates. Thereby, they neglect to explain the continuum between highly derogatory and purely descriptive terms. Understanding this continuum as depending both on differences in inferential commitments and on different degrees of their explicitness, however, is crucial to see that the meaning of terms and their derogatory force may change over time if their underlying social practices change.

6 Conclusion

In this paper, I defended an inferentialist explanation of derogatory terms against the main objections that are usually raised against the inferentialist view. I rejected three of the four main objections by showing that they are based on misunderstandings of the core assumptions of inferentialist semantics. These criticisms can be accommodated by pointing to the role that conceptual norms play in inferentialist semantics, the degrees of conceptual competence, and to the social embeddedness of linguistic and non-linguistic moves that can be made in the inferentialist scorekeeping practice.

doi: 10.48106/dial.v75.i3.05

I further argued both that the indeterminate reference objection rests on a misunderstanding of Dummett's and Brandom's rules for the application and use of derogatory terms, and that this misunderstanding is productive in that these rules are indeed misguided for a broad range of derogatory terms. Dummett and Brandom discuss their "b·che" example as a case in which the inferences that are licensed by the use of the term are not yet part of the web of inferences that were relevant to determine whether the application of the term was appropriate. This usually happens whenever the meaning of a term changes. If the meaning of a term changes, either the rules of application or the consequences of using the term change first. Initially, these changes are only implicit in the practices. By making these changes explicit, we bring the new inferences into the realm of reasoning. Brandom thinks that this mechanism not only underlies meaning change, but also the use of derogatory terms.

However, this is not the case. At most it explains the mechanism involved whenever members of a discursive community who thought that a term was purely descriptive become aware of its derogatory function. Historical examples are "Ch·naman" and the US American use of "Or·ental". In such cases, these members thought that the introduction rule for "*x* is a Ch·naman" were "*x* is a Chinese man" and they learn that the elimination rule is "*x* is less civilised than Europeans". As in the case of meaning change, making explicit the derogatory consequences of a term's use amounts to a change of the application rules, too. In a case of meaning change, for instance, it was no longer possible to have Newton's application rules for "gravity" once Einstein's theory of relativity had been widely adopted. The application rules had to be adjusted to be in accordance with Einstein's theory. Analogously, in the case of explicitly derogatory terms, it is no longer possible to turn to the past ignorant discursive community—which took "Ch·naman" to be a neutrally descriptive term while implicitly treating it as a derogatory one—if we look for the current application rules. Once the pernicious consequences of the term's use are made explicit, its application rules change, too.

Accounting for explicitly derogatory terms thus amounts to including a whole set of inferences into the rules of application, some of which might involve evaluative components. If we do not allow devaluating claims to enter the set of inferences that is relevant to determine whether the application of a term is appropriate, we not only have to face Williamson's and Hom's indeterminate reference objection, but we also misdescribe the practice of derogatory language use. Instead of following Dummett and Brandom in

looking at the commitments of language users who refrain from using a specific term, we need to turn our gaze toward the commitments of those who engage in the practice of using the explicitly derogatory term in question. Highlighting the commitments of bigoted language users, in turn, suggests that a successful critique of derogatory language use does not merely consist in problematising the use of a specific term, but rather taking issue with a whole set of practices in which the use of this term is embedded. These practices may involve a broad range of sets of commitments, from highly pernicious commitments to sets of commitments that are not morally pernicious. The view I have put forward is thus an inferentialist account of conceptual content in general which treats explicitly derogatory concepts as part of a continuum with implicitly derogatory terms and terms that are not derogatory at all.*

Deborah Raika Mühlebach
0000-0002-5720-3437
Freie Universität Berlin
d.muehlebach@fu-berlin.de

References

Ahmed, Sara. 2004. "Declarations of Whiteness: The Non-Performativity of Anti-Racism." *Borderlands* 3(2). Republished as Ahmed (2006).

—. 2006. "The Nonperformativity of Antiracism." *Meridians* 7(1): 104–126, doi:10.2979/mer.2006.7.1.104. Republication of Ahmed (2004).

Alcoff, Linda Martín. 2015. *The Future of Whiteness*. Cambridge: Polity Press.

Anderson, Luvell. 2018. "Calling, Addressing, and Appropriation ." In *Bad Words. Philosophical Perspectives on Slurs*, edited by David Sosa, pp. 6–28. Oxford: Oxford University Press, doi:10.1093/oso/9780198758655.001.0001.

Ashwell, Lauren. 2016. "Gendered Slurs ." *Social Theory and Practice* 42(2): 228–391, doi:10.5840/soctheorpract201642213.

Brandom, Robert B. 1994. *Making It Explicit: Reasoning, Representing, and Discursive Commitment*. Cambridge, Massachusetts: Harvard University Press.

—. 2000. *Articulating Reasons: An Introduction to Inferentialism*. Cambridge, Massachusetts: Harvard University Press.

* Thanks to Sally Haslanger, Samia Hesni, Rebekka Hufendiek, Justin Khoo, Dominique Kuenzle, Michael O'Leary, Melanie Sarzano, Jennifer Saul, Christine Sievers, Pietro Snider, Marie van Loon, and Markus Wild for helpful comments on earlier drafts of this paper. Special thanks to Matthieu Queloz for inspiring discussions about issues raised here. I am grateful to the anonymous reviewers of this and other journals whose comments helped to significantly improve the content of this paper.

—. 2001. “Reason, Expression, and the Philosophic Enterprise.” In *What Is Philosophy?*, edited by C. P. RAGLAND and Sarah HEIDT, pp. 74–95. New Haven, Connecticut: Yale University Press.

CAMP, Elisabeth. 2013. “Slurring Perspectives.” *Analytic Philosophy* 54(3): 330–349, doi:10.1111/phib.12022.

CEPOLLARO, Bianca, and T. THOMMEN. 2019. “What’s Wrong with Truth-Conditional Accounts of Slurs .” *Linguistics and Philosophy* 42(4): 333–347, doi:10.1007/s10988-018-9249-8.

CROOM, Adam M. 2011. “Slurs.” *Language Sciences* 33: 343–358, doi:10.1016/j.langsci.2010.11.005.

DÍAZ-LEÓN, Esa. 2020. “Pejorative Terms and the Semantic Strategy.” *Acta Analytica* 35(1): 23–34, doi:10.1007/s12136-019-00392-2.

DUMMETT, Michael A. E. 1973. *Frege: Philosophy of Language*. London: Gerald Duckworth & Co.

GLEZAKOS, Stavroula. 2013. “Words Gone Sour?” In *Reference and Referring*, edited by William P. KABASENCHE, Michael O’ROURKE, and Matthew H. SLATER, pp. 385–404. Topics in Contemporary Philosophy n. 9. Cambridge, Massachusetts: The MIT Press, doi:10.7551/mitpress/9581.001.0001.

HIGGINBOTHAM, James. 1998. “Conceptual Competence.” In *Philosophical Issues 9: Concepts*, edited by Enrique VILLANUEVA, pp. 149–162. Atascadero, California: Ridgeview Publishing Co., doi:10.2307/1522965.

HOM, Christopher. 2008. “The Semantics of Racial Epiteths.” *The Journal of Philosophy* 105(8): 416–440, doi:10.5840/jphil2008105834.

—. 2010. “Pejoratives.” *Philosophy Compass* 5(2): 164–185, doi:10.1111/j.1747-9991.2009.00274.x.

—. 2012. “A Puzzle about Pejoratives.” *Philosophical Studies* 159(3): 383–405, doi:10.1007/s11098-011-9749-7.

HOM, Christopher, and Robert C. MAY. 2013. “Moral and Semantic Innocence.” *Analytic Philosophy* 54(3): 293–313, doi:10.1111/phib.12020.

—. 2018. “Pejoratives as Fiction .” In *Bad Words. Philosophical Perspectives on Slurs*, edited by David SOSA, pp. 108–131. Oxford: Oxford University Press, doi:10.1093/oso/9780198758655.001.0001.

HORNSBY, Jennifer. 2001. “Meaning and Uselessness: How to Think about Derogatory Words.” In *Midwest Studies in Philosophy 25: Figurative Language*, edited by Peter A. FRENCH and Howard K. WETTSTEIN, pp. 128–141. Boston, Massachusetts: Basil Blackwell Publishers, doi:10.1111/1475-4975.00042.

HYDE, Harford Montgomery. 1962. *The Trials of Oscar Wilde*. Mineola, New York: Dover Publications.

JAMES, Aaron. 2012. *Assholes. A Theory*. Garden City, New York: Doubleday & Co.

JESHION, Robin. 2013. "Expressivism and the Offensiveness of Slurs." In *Philosophical Perspectives 27: Philosophy of Language*, edited by John HAWTHORNE, pp. 231–259. Hoboken, New Jersey: John Wiley and Sons, Inc., doi:10.1111/phpe.12027.

MCCONNELL-GINET, Sally. 2006. "Why Defining Is Seldom 'Just Semantics': Marriage and Marriage." In *Drawing the Boundaries of Meaning. Neo-Gricean Studies in Pragmatics and Semantics in Honor of Laurence R. Horn*, edited by Betty J. BIRNER and Gregory WARD, pp. 217–240. Studies in Language, Companion Series n. 80. Amsterdam: John Benjamins Publishing Co.

—. 2011. *Gender, Sexuality, and Meaning. Linguistic Practice and Politics*. Oxford: Oxford University Press.

MÜHLEBACH, Deborah Raika. 2021. "Semantic Contestations and the Meaning of Politically Significant Terms ." *Inquiry* 64(8): 788–817, doi:10.1080/0020174x.2019.1592702.

—. 2022. "Tackling Verbal Derogation: Linguistic Meaning, Social Meaning and Constructive Contestation." In *The Political Turn in Analytic Philosophy. Reflections on Social Injustice and Oppression*, edited by David BORDONABA-PLOU, Víctor Fernández CASTRO, and José Ramón TORICES, pp. 175–198. Eide n. 11. Berlin: Walter De Gruyter.

—. 2023a. "A Non-Ideal Approach to Slurs." *Synthese* 202(3), doi:10.1007/s11229-023-04315-y.

—. 2023b. "Meaning in Derogatory Social Practices." *Theoria* 89(4), doi:10.1111/theo.12476.

TIRRELL, Lynne. 1999. "Derogatory Terms: Racism, Sexism, and the Inferential Role Theory of Meaning." In *Language and Liberation: Feminism, Philosophy, and Language*, edited by Christina HENDRICKS and Kelly OLIVER, pp. 41–80. New York: State University of New York Press.

—. 2012. "Genocidal Language Games." In *Speech and Harm: Controversies over Free Speech*, edited by Ishani MAITRA and Mary Kate MCGOWAN, pp. 174–221. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780199236282.001.0001.

WHITING, Daniel. 2007. "Inferentialism, Representationalism and Derogatory Words." *International Journal of Philosophical Studies* 15(2): 191–205, doi:10.1080/09672550701383483.

—. 2008. "Conservatives and Racists: Inferential Role Semantics and Pejoratives." *Philosophia* 36(3): 375–388, doi:10.1007/s11406-007-9109-1.

—. 2013. "It's Not What You Said, It's the Way You Said It: Slurs and Conventional Implicatures." *Analytic Philosophy* 54(3): 364–377, doi:10.1111/phib.12024.

WILLIAMSON, Timothy. 2009. "Reference, Inference, and the Semantics of Pejoratives." In *The Philosophy of David Kaplan*, edited by Joseph ALMOG and Paolo LEONARDI, pp. 137–158. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780195367881.001.0001.

YOUNG, Iris Marion. 1990. *Justice and the Politics of Difference*. Princeton, New Jersey: Princeton University Press. Second edition Young (2012).

—. 2012. *Justice and the Politics of Difference*. 2nd ed. Princeton, New Jersey: Princeton University Press, doi:10.1515/9781400839902. With a foreword by Danielle S. Allen.