

Rethinking Slurs: A Case Against Neutral Counterparts and the Introduction of Referential Flexibility


Alice Damirjian*

Received: 4 December 2020 / Revised: April 14 2021 / Accepted: 2 June 2021

Abstract: Slurs are pejorative expressions that derogate individuals or groups on the basis of their gender, race, nationality, religion, sexual orientation and so forth. In the constantly growing literature on slurs, it has become customary to appeal to so-called “neutral counterparts” for explaining the extension and truth-conditional content of slurring terms. More precisely, it is commonly assumed that every slur shares its extension and literal content with a non-evaluative counterpart term. I think this assumption is unwarranted and, in this paper, I shall present two arguments against it. (i) A careful comparison of slurs with complex or thick group-referencing pejoratives lacking neutral counterparts shows that these are in fact very hard to distinguish. (ii) Slurs lack the referential stability of their alleged neutral counterparts, which suggests that they are not coreferential. Developing (ii) will involve introducing a new concept which I regard as essential for understanding how slurs behave in natural language: referential flexibility. I shall support my claims by looking at historical and current ways in which slurs and other pejorative terms are used, and I shall argue that both etymological data and new empirical data support the conclusion that the assumption of neutral counterparts not only is unwarranted but obscures our understanding of what slurs are, and what speakers do with them.

* Stockholm University

 <https://orcid.org/0000-0003-4410-2087>

 Universitetsvägen 10D, 114 18 Stockholm, Sweden.

 alice.damirjian@philosophy.su.se

© The Author. Journal compilation © The Editorial Board, *Organon F*.



This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International Public License (CC BY-NC 4.0).

Keywords: Neutral counterparts, pejorative language, philosophy of language, semantics, slurs.

1. Introduction

In the constantly growing literature on slurs, it has become customary to appeal to so-called “neutral counterparts” for explaining the extension and truth-conditional content of slurring terms. More precisely, it is commonly assumed that every slur shares its extension and literal content with a non-evaluative counterpart term. Paradigmatic examples of slurs with respective neutral counterparts in the literature are: ‘n*gg*r’ with ‘Black person’, ‘k*ke’ with ‘Jew’ and ‘ch*nk’ with ‘Chinese’.¹

This assumption often takes the form of an identity thesis, which states that slurs are extensionally and truth-conditionally equivalent with their relevant neutral counterparts. In his often-cited paper, Whiting summarizes his view in the following way: “[...] what is said by the use of a slur is what is said by the use of its neutral counterpart, whereas what is derogatory about the use of a slur is the claim or, as I prefer, attitude conventionally implicated by it” (2013, 376). In a similar vein, Vallée argues that: “If S is an ethnic slur in language L, then there is a non-derogatory expression G in L such that G and S have the same extension” (2014, 79). It is this suggested equivalence between slurs and their alleged neutral counterparts that will be addressed in this paper.

Seeing as this assumption about the truth-conditional content of slurring terms has become so widely accepted, the debate surrounding slurs has largely come to be about what *other* properties slurs should be ascribed, that can account for their offensive nature. Some take the derogatory content of a slur to be expressed by way of a conventional implicature, so that it becomes part of the slur’s conventional meaning, but not its truth-

¹ CONTENT WARNING: Here, I have chosen to transcribe the slurs with asterisks in order to avoid causing unnecessary harm, and I will continue to do so throughout this paper. However, for clarity, less common slurs have been left unedited, and I will also not transcribe slurs occurring in quotes from other authors or sentence examples taken from social media.

conditional content (cf. Whiting 2013; Williamson 2009). Others prefer to analyze it as a presupposition (cf. Schlenker 2007; Cepollaro 2015), or as a violation of prohibition (Anderson and Lepore 2013a, 2013b). Yet others suggest that we must focus on what speakers are doing when they *choose* to use a slur instead of its non-derogatory counterpart term. Examples of such accounts are Camp's (2013, 2018) perspectivist account and pragmatic accounts like Bolinger's (2017) and Nunberg's (2018).² To illustrate, Ashwell has described Camp as giving us

[...] the most persuasive reason for thinking that there has to be a neutral alternative term for a word to be a slur – that it is only in contrast with an alternative neutral term that the bigot can signal their bigoted perspective. (Ashwell 2016, 238)

Of course, how much a specific account relies on the existence of neutral counterparts varies. Prohibitionist and conventional implicature views are, for example, not necessarily reliant on neutral counterparts.³ However, many accounts require the existence of the neutral counterpart for their preferred analysis to work, and this is especially pressing for accounts that place weight on the *optionality* of slurs, e.g. Bolinger's and Nunberg's.⁴ But if the assumption of neutral counterparts turns out to be non-viable, all proponents of accounts that rely on them to some degree will have to re-think how they explain the truth-conditional content of slurring terms.

In this paper, I shall argue that the pervasive idea that neutral counterparts can play this semantic role in our theories of slurs is unmotivated and that it obscures our understanding of what slurs are and what speakers do with them. This will be supported by a body of evidence that casts serious doubt upon the assumption that the extension of a slur is identical to the extension of some salient neutral counterpart. Then, if we are not warranted in assuming that every slur is extensionally equivalent with a neutral counterpart, we cannot assume that every slur shares its literal meaning with

² Undeniably, this list is very far from complete.

³ Since it in principle could be argued that a slur is prohibited, or has the relevant conventional implicature, without it being extensionally or truth-conditionally equivalent with a neutral counterpart.

⁴ See Falbo (2021) for a more detailed explanation for why this is.

one either. The central claim is thus that the identity thesis is deeply flawed, and the important consequence of this is that we must rethink how we analyze the semantics of slurs, so as to be able to account for and predict the data that will be provided here. The paper will not be proposing an alternative positive account of the meaning of slurs; instead, the aim is to provide insights needed for moving forward in the field.

2. Neutral counterpart theories

Diaz Legaspe (2018) distinguishes between two ways in which theories of slurs rely on neutral counterparts. Firstly, we can identify a weaker position, widely accepted in most of the literature, which Diaz Legaspe calls the *Application Neutral Counterpart Thesis* (AT). Secondly, there is a stronger position not as widely accepted, called the *Referential Neutral Counterpart Thesis* (RT). The two theses are defined as follows (Diaz Legaspe 2018, 235):

AT: For every slurring expression e there is a neutral counterpart N_{Ce} such that N_{Ce} 's correct application criteria are identical to e 's correct application criteria.⁵

RT: For every slurring expression e there is a neutral counterpart N_{Ce} such that the class of individuals referred to by N_{Ce} (call it $\{N_{Ce}\}$) is identical to the class of individuals referred to by e .⁶

AT is intended to describe accounts similar to Hom's (2010, 2012) and Hom and May's (2013), which pack the derogatory content of slurs into their

⁵ Diaz Legaspe argues that the N_{Ce} can be an actual or potential neutral counterpart to e , "potential" meaning that the neutral counterpart might not have been found yet or that it is for some reason unavailable. This, however, strikes me as strange. Either there already exists a neutral counterpart to e in the language or e lacks a neutral counterpart, we should not just be able to state that e could be given a neutral counterpart.

⁶ The notation might strike the reader as odd. To clarify, N_{Ce} is not a set, but an abbreviation for the "neutral counterpart to e ." $\{N_{Ce}\}$ is denoting a set: the class of individuals referred to by N_{Ce} . With that said, I will continue to use these notations as introduced by Diaz Legaspe.

semantics. On Hom's account, every slur targets some specific group picked out by its neutral counterpart, but the slur and the neutral counterpart are still taken to diverge in meaning and extension (in fact, the extension of a slur is assumed to be the empty set). Hence, AT is silent regarding the truth-conditional contribution of a slur. "Correct application" is therefore not intended to be understood in a strong sense.⁷ Instead, "correct application" simply makes it a necessary condition that the individual referred to by *e* is a member of {N*Ce*} for *e* to have been correctly applied. So, the thesis does predict that it is a linguistic mistake to use *e* for someone outside of {N*Ce*}, but it does not claim *e* and N*Ce* to be intersubstitutable *salva veritate*. In contrast, RT states that the extension of a slur will be the same as that of its neutral counterpart, and this is supposed to be a consequence of the terms' making the same truth-conditional contribution.⁸

RT has turned out to be a common approach for explaining the semantics of slurring terms, and this is standardly the case for the views included here in §1. Consequently, it is primarily the identity thesis at the heart of RT which this paper will target. Even so, since AT does predict that it is a linguistic mistake to use *e* for an individual outside of {N*Ce*}, AT will also be sensitive to the arguments presented here.

3. Neutral counterpart skepticism

There have been attempts to show that the whole idea of co-referentiality between a slur and a non-pejorative synonym is misguided. For example, Croom (2015) has argued that empirical data shows that slurs cannot be coreferential with a neutral counterpart. Croom provides the reader with

⁷ For a strong reading would suggest co-referentiality. AT is intended to describe and provide "a conception of correct application that does not amount to truth-conditional identity" (Diaz Legaspe 2018, footnote 5, 135).

⁸ This is not explicitly stated in RT but it is assumed that co-referentiality amounts to truth-conditional equivalence. "In turn, RT predicts that for every slur and every associated neutral counterpart, both will contribute the same set of individuals to the truth-conditions of the utterance in which they occur" (Diaz Legaspe 2018, 235).

four examples, taken from academic research, literature and comedy, of when speakers are employing slurs in ways inconsistent with what RT predicts.⁹ These examples are, even if Croom does not call them that, examples of *referential restriction* (a phenomenon we will come back to in §4.2); cases of slurring in which the slur is used not to pick out all of {NCe} but only a subset. For example, when ‘n*gg*r’ is applied to some *but not all* Black people, or when ‘f*ggot’ is deemed not to apply to all gay men, but only to *some* male homosexuals. This leads Croom to the conclusion that

[...] the fact that the slur *faggot* is differentially used so that it is often applied to *some but not all* male homosexuals suggests that the slur *faggot* and the descriptor *male homosexual* are in fact *not coreferential expressions with precisely the same extension* at all. (2015, 32)

If slurs are not coreferential with a neutral counterpart, then a slur cannot be said to share its meaning with one either, Croom argues, and any theory of slurs will have to respect that fact. This is an observation that will be of importance throughout this paper.

Croom is not the only one to criticize the identity thesis at issue here. Ashwell (2016) has argued that all accounts of slurs that assume neutral counterparts will fail to generalize over all terms we are inclined to regard as slurs. According to Ashwell this is best shown when considering gendered slurs, such as ‘slut’ or ‘bitch’, which have not been given the same attention as racial or ethnic slurs in the literature.

Some might want to argue (e.g., Diaz Legaspe 2018) that ‘bitch’ and similar gendered slurs have ‘woman’ as their neutral counterpart, but it appears unlikely that bigoted individuals would hold that “*all* women are bitches” in the same way as they would hold that “*all* Jews are k*kes” or “*all* Blacks are n*gg*rs.” Of course, ‘bitch’ and ‘slut’ are generally applied to women, but one is not a ‘slut’ for being a woman. One is a ‘slut’ in virtue of something else, Ashwell claims (2016, 234), something that has to do with sex. However, “woman who has sex with a lot of partners” cannot function as a neutral counterpart, it is not free of pejorative association nor

⁹ The examples will not be presented here but can be found in (Croom 2015, 32-34).

is it neutral because “a lot of partners” is dependent on what is assumed to be the appropriate number of sexual partners (Ashwell 2016, 234-35).

Are gendered slurs actually slurs then, or some other kind of pejorative? “They are not slurs” is the answer scholars like Nunberg (2018) would give: they are what is called *hybrid words* and therefore not slurs in the strict sense (see §4.1 in this paper). Objections like Ashwell’s can therefore be dismissed by claiming that the terms brought forward as counterexamples are not in fact the kind of pejoratives that we are interested in; not slurs but some other type of pejorative. However, if one wants to go down that road one must be careful, as DiFranco (2015) has pointed out:

They [those assuming neutral counterparts] should not simply insist that the class of conventional slurring words is, by definition, restricted to words and phrases whose truth-conditional content is identical to that of their neutral counterparts. Doing so would beg the question by illicitly presupposing NC [truth-conditional equivalence]. (2015, 33)

One should not simply claim that the class of slurs is exactly that class of non-complex slurring terms that share their truth-conditional content with a neutral counterpart. Such a claim would have to be motivated. It could be motivated if: (i) It can be shown that the class of slurs is significantly large and clearly distinguished from other kinds of pejoratives lacking neutral counterparts; or (ii) we have strong evidence that slurs express the same thing, on a truth-conditional level, as their alleged neutral counterparts. For the remainder of this paper, I shall argue that no such motivation exists.

4. Against neutral counterparts

In the following sections I will argue for two points, each of which, and especially in combination, should be regarded as severely undermining the credibility of appealing to the identity thesis in one’s theory of slurs. The first point has to do with the characterization of slurs, and I will argue that we cannot distinguish a distinct class of slurs from other pejoratives lacking neutral counterparts. The second point is that slurs can be used to refer in

flexible ways, a feature of slurs that their alleged neutral counterparts lack. As argued by Croom, speakers can use slurs with referential restriction to refer to a subset of the set {N*C*e}, but I will argue that speakers also use slurs to refer outside of {N*C*e}, in ways that cannot be disregarded as linguistic mistakes. I will call this phenomenon *referential expansion*. This will lead us to the conclusion that slurs are used, and can be used, in flexible ways, both restrictedly and expansively. I will call this feature *referential flexibility*.¹⁰

Referential flexibility can be observed both historically and in present-day use, and I will therefore spend a fair amount of time on etymological examples and speaker examples that I have found on the internet. Already, I would like to make clear that the point of the etymological examples is not say that slurring terms somehow have kept the meaning they have had historically. Rather, it is to show that their application criteria have been, and still are, flexible – or at least not non-flexible in the sense RT and AT suggest.

4.1 *Distinguishing the non-distinct distinctions*

Hybrid words and *umbrella derogatives* are terms usually regarded as slurs, but which appear to lack neutral counterparts.¹¹ Hybrid words, as

¹⁰ It should be noted that once referential flexibility has been introduced, talk of referential expansion will soon prove problematic, for such talk implies the existence of a neutral counterpart that can be expanded upon. The reader should therefore keep in mind that the characterization of referential expansion as it is introduced here is not intended to serve any explanatory purposes apart from the argumentative role it plays within the scope of this paper. Once we have dropped neutral counterparts, there is no further need to discuss the kind of referential expansion I introduce. Hopefully, all of this will become clear after §4.2.

¹¹ The discussion here will primarily focus on umbrella derogatives, as they pose the greatest challenge to theories relying on neutral counterparts. In principle, hybrid words could be dealt with by holding that the additional evaluative content they incorporate does not contribute to determining the terms' extensions. This way out is suggested by Jeshion (2013, 234-235). However, for scholars like Nunberg (2018) who hold that hybrid words should not be conflated with slurs, the discussion here about hybrid words is highly relevant.

defined by Nunberg (2018), are terms that do not only categorize but mix categorization and attitude, so that some strong evaluation of the referent is part of the content of the term that cannot be found in any non-evaluative synonym. They are sometimes referred to as “thick terms”, precisely because they incorporate an evaluation or a stereotypical content within their semantics. Examples are ‘wetback’ for (illegal) Mexican migrants, ‘JAP’ (‘Jewish American Princess’) for spoiled Jewish women, and ‘Uncle Tom’ for Black people who behave in subservient ways towards white people. ‘Slut’, ‘bitch’ and most other disparaging terms for women are also taken to fall into this category (see Nunberg 2018, 249-250).¹²

Umbrella derogatives are pejoratives for collections of distinct groups which become problematically grouped together, to the extent that the terms can have no neutral counterpart.¹³ Clear-cut examples are ‘slope’, mostly used for East Asians, ‘wog’ standardly used for any non-white foreigner, ‘dago’ sweepingly used for Italians, Spaniards and Portuguese, and ‘gook’ for foreigners, especially those of East Asian descent.¹⁴ There rarely exists a synonym for referring to the same group, the group itself is hard to distinguish, and even if we were to stipulate a potential neutral counterpart it would prove difficult to find a non-evaluative one.¹⁵ For that reason,

¹² Whether or not we are dealing with something similar to thick ethical concepts, such as ‘brave’ and ‘generous’, can be debated, but nothing in this discussion hinges on whether such a similarity exists. For the purposes of this paper, it is enough to observe that there is a structural similarity insofar as both hybrid words and standard examples of thick terms exhibit this sort of hybrid nature. I thank the anonymous referee who drew my attention to this unclarity.

¹³ Note that this does not entail that umbrella derogatives lack semantic meaning, just that there does not exist any salient neutral counterpart with which they could be said to share literal meaning.

¹⁴ Jeshion (2016, 135) also notes that slurs like ‘wop’, ‘dago’ and ‘gook’ fail to possess neutral counterparts.

¹⁵ Why cannot, for example, ‘non-white foreigner’ play the neutral counterpart-role? The problem is not that this description does not pick out a category that it could be said to refer to, the problem is that ‘wog’ does not mean ‘non-white foreigner’. When I say that ‘wog’ is used for non-white foreigners it is a simplified account of how speakers tend to use the term, not that it is coreferential with ‘non-white foreigner’. Furthermore, ‘non-white foreigner’ provides us with a large collection of

a theory of slurs that places weight on neutral counterparts in its analysis will fail to deal with umbrella derogatives.

According to Nunberg (2018), these terms should not be confused with slurs, which in a stricter sense belong to the class of pejoratives that have a non-slurring counterpart. This is important for his account to work, and for many others. However, is it reasonable to think that such a class of slurs can be said to exist in any well-defined sense, such that slurs, as a well-considered category, can be singled out as Nunberg suggests?

When looking closer, we can observe interesting similarities between umbrella derogatives like ‘gook’ and pejoratives like ‘ch*nk’ and ‘j*p’, both perceived as unproblematic examples of slurs. ‘Gook’ stands out in that the term has come to play many different xenophobic roles in a relatively short period of time. As a pejorative, it has not only been used against people of East Asian descent but for any foreigner (from an American perspective) and any foreign language (i.e., not English).¹⁶ In his encyclopedia of swearing, Hughes explains that “[...] its semantic history combines hostility toward outsiders with great flexibility in application” (2006, 207). But ‘gook’ is not the only term that has played varying roles; ‘ch*nk’ does in fact have a similar, but not as striking, history of flexibility in application.

During the 1849 California Gold Rush a great number of Chinese immigrants arrived in America to work as indentured laborers, and the resentment towards the immigration of cheap labor within the group of white native-born laborers resulted in several names for the out-group. The dominant derogative for the Chinese immigrants was ‘ch*nk’ (Hughes 2006, 75-

distinct groups that should not be conflated and so arbitrarily referred to as a single group. Thus, even if we stipulated that ‘non-white foreigner’ gives us the meaning of ‘wog’, it would not be neutral in the sense required. Imagine a speaker uttering ‘There were so many k*kes in the park today.’ The (supposedly) equivalent utterance ‘There were so many Jews in the park today’ can reasonably be said to be neutral in the sense required (if there is nothing in the context of utterance to suggest otherwise). Arguably, the same is not the case for ‘There were so many non-white foreigners in the park today.’ Standardly, this is taken to be what distinguishes slurs from umbrella derogatives.

¹⁶ The etymological data presented here, and in other places in the paper, comes from Green’s *Chambers Slang Dictionary* (2008) and Hughes’ *An Encyclopedia of Swearing* (2006).

76). The term was used for Chinese people but also for people with Chinese features, and thus more generally for any person of East Asian descent. It was especially from 1942 that ‘ch*nk’ began to be applied to any East Asian person, which coincides with the evolution of ‘j*p’. The abbreviation ‘j*p’ for ‘Japanese’ was common from around the 1850s, but not necessarily offensive and not exclusively used for Japanese. But after Pearl Harbor, terms of abuse arose rapidly and ‘j*p’ was simultaneously used for Japanese people and as a slang for being sneaky or a bad surprise. Other people with similar appearances were conflated too, others from the Far East, all of whom were labeled ‘gooks’ (Hughes 2006, 262).

So, are these slurs better described as umbrella derogatives rather than actual slurs in the relevant sense? The answer depends on how speakers today use the slurs, which is an empirical question, but it is plausible to assume that slurs are used in this loose way quite often. Consider this testimony from a man of Cambodian descent:

This guy in DC just skipped me in line at 7/11 [7-Eleven] and then proceeded to call me a ch*nk – multiple times. I told him he had skipped me in line and that’s when he got aggressive. Anti-Asian racism is real and it’s fucked up. I’m okay, just a jarring experience.¹⁷

The example illustrates two important points: (i) That the speaker deems it relevant to utilize the term ‘ch*nk’, not knowing or not caring about the actual nationality of the targeted person, and (ii) that the targeted person himself makes a point of saying that the speaker’s words were anti-Asian, rather than anti-Chinese.

Additionally, one could also argue that ‘j*p’ functions more like a hybrid word, i.e., that one can observe it possessing the properties deemed distinctive of a hybrid word. Nunberg claims that hybrid words are distinguished from slurs by the fact that the evaluative content present in hybrid words can be contested, but not restated without a feeling of redundancy. In contrast to actual slurs, hybrid words carry their evaluative content within

¹⁷ <https://tinyurl.com/tjsb5mc>, accessed 13 April 2020. All URLs have been shortened as they are sometimes very long, but will redirect the reader to where the examples can be found.

their conventional meaning, and thus a sense of redundancy will arise in sentences like:

- (1) Uncle Toms are really obsequious.
- (2) Bitches are malicious women.

Sentences like (1) and (2) should rightly elicit the reaction “So what else is new?” Nunberg argues (2018, 249); it is already part of the meaning of calling somebody an ‘Uncle Tom’ that they are obsequious or subservient. Yet, with actual slurs that feeling of redundancy does not occur, hence sentences (3) and (4) should appear informative, and indeed common.

- (3) N*gg*rs are so lazy.
- (4) I don’t like that k*ke, he’s very greedy.

Primarily, this is one of Nunberg’s arguments against conventional implicature views, for if the derogatory content is part of the conventional meaning of slurs, then this redundancy should be present in any construction similar to (1)–(4), and not only those involving hybrid words. But it also provides us with a clear description of hybrid words.

Still, it does not seem all that easy to weed out the hybrids from the slurs. Turning back to ‘j*p’, the rapid expansion of the use of ‘j*p’ during the Second World War as, simultaneously, a word for the Japanese soldiers and the Japanese living in the United States (and indeed anyone East Asian looking), and as a slang term for sneaky things and bad surprises, might suggest that the term could be treated as a hybrid word. There is reason to think that the current use of ‘j*p’ carries that evaluative content within its meaning, conventionally, so that one should react with “So what else is new?” in response to (5).

- (5) The j*ps are so sneaky.

This is not necessarily to say that ‘j*p’ is a hybrid word. The point is to illustrate that the distinction between slurs, hybrid words and umbrella derogatives is not crystal clear – in fact, it is not clear at all.

Depending on how you approach the terms, you might get different intuitions about how they are used. If you choose to only study the cases in which ‘ch*nk’ is used for Chinese people, trying to figure out what the term means, you will probably feel that ‘ch*nk’ refers only to Chinese people.

But if you instead start to look at all the cases in which ‘ch*nk’ and ‘j*p’ are used more broadly, you will probably agree that these cases attest to the terms functioning more like umbrella derogatives or hybrid words.

4.2 Arguments for referential flexibility

Diaz Legaspe (2018) attempts to tackle the problematic aspects of referential restriction, which were illustrated in Croom’s argument above, and tries to defend both AT and RT against it. To repeat: referential restriction occurs when a slur is used to refer to a subset of {N_{Ce}}. A commonly used example of the phenomenon is the statement (6) made by comedian Chris Rock.

(6) I love Black people, but I hate niggers.¹⁸

Referential restriction poses a problem for theories relying on RT, Diaz Legaspe argues, because if slurs always share their truth-conditional content with their associated neutral counterparts, then sentences like (6) should be contradictory. However, sentences like (6) are common and appear informative, so some slurs do seem to be able to refer to a narrower class than {N_{Ce}}. Further, Diaz Legaspe observes, some slurs even appear to always refer to a sub-class of their neutral counterpart, such as gendered slurs.

To try and solve this, Diaz Legaspe proposes modifications to AT and RT in order to restore the link between slurs and their neutral counterparts. Diaz Legaspe’s proposal is, roughly, that there can be particular contexts in which a slur *e* (e.g. a racial slur), whose reference is {N_{Ce}}, can be used to refer to a subset of {N_{Ce}}, but there are also some slurs (e.g. all gendered slurs) which *always* refer to a subset of the class picked out by their neutral counterparts.¹⁹ Thus, racial slurs and gendered slurs should be understood differently, but even so, gendered slurs can be assigned neutral counterparts (holding that the slurs just always refer to a subset of the set picked out by

¹⁸ The example appears in (Diaz Legaspe 2018, 236 and 243). This is also one of Croom’s examples (2015, 33) and it appears in: (Nunberg 2018, footnote 12, 247), (Rappaport 2019, 810), (Jeshion 2013, 233 and 238-239), (Anderson and Lepore 2013b, footnote 3, 43), to name a few occurrences.

¹⁹ This approach is partly aimed at solving the problems with gendered slurs set up by Ashwell (2016).

their neutral counterparts). These are the modified versions of the two theses, which are assumed to hold whether you accept RT and AT or just AT, and for all types of slurs (Diaz Legaspe 2018, 248-249):

Negative AT: For any slur e there is a NCe such that every member of $\{\text{NCe}\}$ can be correctly called an ‘ e ’.

Furthermore, *only* members of $\{\text{NCe}\}$ can correctly be called an ‘ e ’, with the exception of metaphorical uses (Diaz Legaspe 2018, 248).²⁰ To call someone outside of $\{\text{NCe}\}$ an ‘ e ’, will count as a linguistic mistake. The second condition is:

*Positive AT**: For every e there is a NCe such that every member of $\{\text{NCe}\}$ could potentially be called an ‘ e ’.²¹

That is, all in $\{\text{NCe}\}$ can potentially be called an ‘ e ’, without it amounting to a linguistic mistake. Seeing as some slurs, like gendered slurs, always refer to a subset of $\{\text{NCe}\}$ “normal” RT will not hold for them, Diaz Legaspe argues, but if one generally wants RT to hold for slurs one can appeal to *Restricted RT*:

Restricted RT: Whenever “ o is an e ” is true, “ o is a NCe ” is also true.

When we are dealing with gendered slurs, however, e and NCe will not be interchangeable in the other direction.

Now, if a case of referential restriction is a case in which a slur e is used to refer to a subset of $\{\text{NCe}\}$, as in sentence (6), then referential expansion is a case in which e is used to refer outside of $\{\text{NCe}\}$, in a way that cannot be disregarded as a linguistic mistake. Of course, such expansions are not allowed once one has accepted the identity thesis, and thus cases of refer-

²⁰ After having introduced the condition, Diaz Legaspe writes “[...] only women can be correctly called ‘sluts’ [...]” (2018, 248).

²¹ This condition (Positive AT*) is a modified version of Positive AT:

Positive AT: For every e there is a NCe such that every member of $\{\text{NCe}\}$ can be called an ‘ e ’.

A condition which Diaz Legaspe claims not to hold, since it is assumed that some slurs always refer to a subset of $\{\text{NCe}\}$. But, she concludes, potentially everyone in $\{\text{NCe}\}$ could be called ‘ e ’.

ential expansion should rightly be treated as counterexamples to the identity thesis. To be clear, the referential expansion that interests us here does not occur from referencing outside of {NCe} in any intentionally incorrect or figurative way by alluding to stereotypes. Neither is referential expansion in our sense simply some linguistic effect a speaker can evoke by referencing outside of {NCe}.

Slurs can expand and gain flexibility in different ways. One way, which is discussed in Hughes' encyclopedia, is when a slur develops from its basic noun function to be used as an adjective or a verb. Standardly, this is seen as an indicator of the slur having become very assimilated into the language (Hughes 2006, 149). However, this is not the kind of expansion that interests us either.²² What interests us is when slurs are actually used to refer to individuals outside of the set {NCe}. The example which supported treating 'ch*nk' as an umbrella derogative is an example of referential expansion.

A category of slurring terms not generally discussed are derogatory terms for disabled people, which are interesting because, like gendered slurs, they too seem difficult to capture within the frameworks of preexisting theories. Such slurs also exhibit the phenomenon of referential expansion because they are commonly used expansively in several different ways. Consider for example the slurring term 'spastic', and alterations such as 'spaz' and 'spazzie', and how they have come to be used. 'Spastic' was first used as a non-derogatory term for people with cerebral palsy, subject to muscle spasm or spasticity, but also became a pejorative for that same group. Understanding 'spastic' as a slur, we should be able to find a salient neutral counterpart in the language. Arguably, its neutral counterpart should be, if

²² The possibility of such morphological transformations could however be seen as supporting my claims. Consider a case in which the noun 'k*ke' is transformed into a verb so that sentences like "He kiked his way to the job" become possible. Or a case of someone uttering "He pulled a jap," or perhaps (a phrase common on the internet) "He's so spazzy." Arguably, what the speaker is intending to convey in cases like the above has little to do with predicating group membership. Rather some descriptive value is intended, and this might suggest that some descriptive, evaluative content is indeed part of the slurs' conventional meaning, also when they are used as nouns. However, it is too complicated a matter for me to be able, within the scope of this paper, to say anything more about it than that it might support my claims.

we were to assign one, ‘people with cerebral palsy’ or ‘people subject to spastic paralysis.’ However, this is not solely how the term tends to be utilized.

Two types of extended uses can be observed, one is an example of what Jeshion calls *G-extending* uses (2013, 238); when speakers use ‘spastic’ of someone *they believe not to have* any physical disability but whom the speaker *wants* to ascribe stereotypical properties associated with people with cerebral palsy, such as being jumpy, clumsy, incapable or incompetent. This kind of expanded use does not necessarily pose a problem for theories postulating neutral counterparts, and therefore it is not this kind of use that interests us.²³ The second kind is when the slur is used for people with other disabilities (people that do not have cerebral palsy), i.e. more broadly of any person with a disability, to derogate them in virtue of that. This can include people with similar symptoms, such as seizures, or other diagnoses associated with similar behaviors, e.g. people living with Tourette syndrome or ADHD. In such cases, we have much less reason to suspect that the relevant speaker is intending to say something that is literally incorrect for some type of linguistic effect. The more reasonable explanation for why this is possible is that the extension of ‘spastic’ is flexible to such a degree that it is not restricted to people with cerebral palsy. That explains why speakers can use ‘spastic’ in this broad way. To exemplify, responding to a post discussing how to calm people with ADHD and overactive children, a Reddit user writes:

(7) Lol just read a book u spastic [sic].²⁴

Diaz Legaspe’s (2018) modified versions of AT and RT were shaped to be compatible with, and even explain, the existence of referential restriction but they are not compatible with referential expansion (nor referential flexibility). If we assume that slurs have the same neutral counterparts as they have been ascribed in previous literature, then every case of referential expansion – every case in which a speaker refers outside of {NCe} with e –

²³ It does not seem problematic to state that this kind of use involves an intentionally incorrect statement, and therefore it does not call into question the appropriateness of the neutral counterpart assumed.

²⁴ <https://tinyurl.com/vfasc19>, accessed 13 April 2020.

should amount to linguistic mistakes following *Negative AT*. *Restricted RT* will also not hold, because if an utterance of “o is an e” is a case of referential expansion it will not also be true that “o is a NCe.” Only *Positive AT** (or even the unmodified *Positive AT*) will still hold, for even if individuals outside of {NCe} can be called an ‘e’, it would still be true that all in {NCe} can potentially be called an ‘e’, but in isolation this condition is very weak. That would open the possibility for many sets to play the neutral counterpart-role, and we would have no way of determining between them, and if it is underdetermined between a number of sets, and potentially all of them could play the role, then referential flexibility is what we get.

Moreover, these kinds of observations cannot be restricted to specific types of slurs, as some might want to claim. Even ‘n*gg*r’ shows an interesting history of flexibility in application. From about the 17th century ‘n*gge*r’ was used to refer to Black people, but primarily slaves. It then evolved to be used for any non-white, around the 19th century, and even more generally for any foreigner, and it also began being used for the Aborigines in Australia.²⁵ Expanded uses, outside of the alleged {NCe} for ‘n*gg*r’, that is ‘Black person’ or ‘African American’, thus seem to have been present for a long time. This tendency is also what allows for expansions with additional content prevalent today, such as ‘sand n*gg*r’ or ‘dune n*gg*r’ for people of Middle Eastern descent, ‘curry n*gg*r’ for Indians, as well as ‘bush n*gg*r’ used for Native Americans, Africans and Aborigines. These are terms that might seem important for the speakers to distinguish, and please excuse the phrase, *just what kind of ‘n*gg*r’ it is they are referring to.*²⁶

²⁵ See (Green 2008, 914-15), for a more detailed account of how the meaning of ‘n*gg*r’ has changed over time.

²⁶ On a side note, there is an interesting question here for compositional semantics about the function of the modifiers ‘sand’, ‘dune’, ‘red’ etc. What are the adjectives actually doing? Normally, an adjective functions so to restrict the extension of the noun it modifies, so that when ‘ball’ is modified with ‘red’ to form ‘red ball’ one has restricted the denotation of ‘ball’ to only include those balls that are red. That is however not what is happening here, if one maintains that ‘n*gg*r’ denotes Black people then ‘red n*gg*r’ cannot be said to restrict the noun (given that ‘red n*gg*r’ is not, standardly, used for Black people at all). Is the modifier then completely changing the extension of the noun? Does it take us outside of the extension of

This can be exemplified with an interaction described by a man working at a hotel in Canada, who describes himself as originally from India. Another man, upset by what he regards as bad customer service, proceeds to call the man ‘sand n*gg*r’ and threatens him as in (8).

- (8) You fucking sand nigger, do you want me to call my boys and have a picnic at your hotel, you fucking piece of shit.²⁷

When asked to leave he continues:

- (9) You fucking nigger, go back to your country, you asshole.²⁸

Sentences (8) and (9) not only illustrate that the speaker in this situation deems it relevant to use ‘n*gg*r’ and ‘sand n*gg*r’ against an Indian man, but also that he regards ‘sand n*gg*r’ as interchangeable with ‘n*gg*r’.

In that sense ‘n*gg*r’ is similar to ‘coon’. In American and British English ‘coon’ has generally been used as a derogatory term for Black people, but not exclusively. Rather, ‘coon’ is similar to ‘wog’ since it has been used more generally for any person of color. It is therefore not surprising that we can see a tendency among speakers to use ‘dune coon’ interchangeably with ‘sand n*gg*r’ and ‘dune n*gg*r’ to refer to Middle Eastern people.

In another Reddit post, a person contemplates the racial slurs that others have called them, which mainly have been words targeting Middle Easterners, such as ‘sand n*gg*r’ and ‘osama’, as well as ‘curry n*gg*r’. Then, in the same post, they observe that:

- (10) Seriously, you can totally change the direction of the n-word just by tacking on a certain word in front of it.²⁹

‘n*gg*r’, in similar ways as ‘fake’ could be said to do when used as a modifier? That might be a possible approach for those assuming neutral counterparts, but it would appear ad hoc to claim that ‘red’ has the capacity to modify ‘ball’ and ‘n*gg*r’ in two very different ways. But, if we allow for the meaning of ‘n*gg*r’ to be broad and flexible like an umbrella derogative’s, we might be able to keep some version of the normal understanding of what the adjectives do.

²⁷ <https://tinyurl.com/ukdvewz>, accessed 13 April 2020.

²⁸ <https://tinyurl.com/ukdvewz>, accessed 13 April 2020.

²⁹ <https://tinyurl.com/vtccllx>, accessed 13 April 2020.

These uses of ‘n*gg*r’ are not G-extending uses nor are they changing the term itself, but they help change the term’s *direction*, that is, its extension. They specify, as stated above, *just what kind of ‘n*gg*r’ it is they are referring to* – which is completely in line with understanding slurs like ‘n*gg*r’ as referentially flexible.

Now, some might want to object and say that my arguments against neutral counterparts will render slurs radically flexible. It was, for instance, pointed out to me by an anonymous referee that referential flexibility will render slurs extensionally unlimited. I very much agree that this is a worry. Of course, there must be some constraints on how slurs can be applied; ethnic slurs derogate groups on the basis of their ethnicity, and it would not seem right (linguistically) to call someone a ‘n*gg*r’ because of their sexual orientation. Even so, understanding that slurs are flexible, to such a degree that the extension of ‘n*gg*r’ is not restricted to any well-defined group, will help explain why they are so difficult to account for philosophically. Referential flexibility is not the end of the story, it invites us to re-think how the meaning of slurs can be accounted for, and my conviction is that when we have found ways of reconceptualizing the meaning of slurs, we will be in a better position to explain why slurs are offensive.

In a recent paper, Falbo (2021) has argued that we must be cautious about assuming that neutral counterparts can play any fundamental or systematic role in explaining the offensiveness of slurs, for in a range of examples neutral counterparts seem unable to do the job they were supposed to do. This observation is right, but we must also accept the stronger claim that neutral counterparts seem unable to play any fundamental role in explaining the truth-conditional content of slurs. Moving forward, we must begin by understanding that racial slurs are used by racists, and *what racists do* is to group people together in arbitrary and insensitive ways – and slurs are a medium for doing just that.

5. Conclusion

To answer the questions leading up to §4: (i) Is the class of slurs significantly large and clearly distinguished from other kinds of pejoratives lacking neutral counterparts? The answer is *no*, many of our most discussed

slurs, such as ‘ch*nk’ and ‘n*gg*r’, are in fact hard to distinguish from umbrella derogatives, and others, like ‘j*p’, are not sufficiently different from hybrid words. Even if we could find some slurs that escape the arguments presented here, they would be rare and therefore any theory of slurs that is only capable of accounting for those few will be insufficiently general. (ii) Do we have strong evidence that slurs express the same thing, on a truth-conditional level, as their alleged neutral counterparts? The answer to this question is also *no*. Since we have evidence for referential flexibility both AT and RT are in trouble, because if slurs and neutral counterparts have non-identical extensions, then they cannot be truth-conditionally equivalent. The conclusion of this paper is thus that the assumption of neutral counterparts is problematic, and that the identity thesis lacks relevant motivations.

Acknowledgments

I would like to thank Dan Zeman and the reviewers for their encouragement and constructive comments, my supervisors Anders Schoubye and Kathrin Glüer-Pagin for their support, and Jesper Olsson for his positivity and sharp eyes.

References

- Anderson, Luvell, and Ernie Lepore. 2013a. “What Did You Call Me? Slurs as Prohibited Words.” *Analytic Philosophy* 54 (3): 350–63.
<https://doi.org/10.1111/phib.12023>
- Anderson, Luvell, and Ernie Lepore. 2013b. “Slurring Words.” *Nous* 47 (1): 25–48.
<https://doi.org/10.1111/j.1468-0068.2010.00820.x>
- Ashwell, Lauren. 2016. “Gendered Slurs.” *Social Theory and Practice* 42 (2): 228–39. <https://doi.org/10.5840/soctheorpract201642213>
- Bolinger, Renée. 2017. “The Pragmatics of Slurs.” *Nous* 51 (3): 439–62.
<https://doi.org/10.1111/nous.12090>
- Camp, Elisabeth. 2013. “Slurring Perspectives.” *Analytic Philosophy* 54 (3): 330–49. <https://doi.org/10.1111/phib.12022>
- Camp, Elisabeth. 2018. “A Dual Act Analysis of Slurs.” In *Bad Words: Philosophical Perspectives on Slurs* edited by David Sosa, 29–59. Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780198758655.001.0001>

- Cepollaro, Bianca. 2015. "In Defence of a Presuppositional Account of Slurs." *Language Sciences* (52): 36–45. <https://doi.org/10.1016/j.langsci.2014.11.004>
- Croom, Adam M. 2015. "The Semantics of Slurs: A Refutation of Coreferentialism." *Ampersand* (2): 30–8. <https://doi.org/10.1016/j.amper.2015.01.001>
- Diaz Legaspe, Justina. 2018. "Normalizing Slurs and Out-Group Slurs: The Case of Referential Restriction." *Analytic Philosophy* 59 (2): 234–55. <https://doi.org/10.1111/phib.12129>
- DiFranco, Ralph. 2015. "Do Racists Speak Truly? On the Truth-Conditional Content of Slurs." *Thought: A Journal of Philosophy* 4 (1): 28–37. <https://doi.org/10.1002/tht3.154>
- Falbo, Arianna. (2021). "Slurs, Neutral Counterparts, and What You Could Have Said." *Analytic Philosophy* (not yet included in an issue). <https://doi.org/10.1111/phib.12217>
- Green, Jonathon. 2008. *Chambers Slang Dictionary*. Edinburgh: Chambers Harrap Publishers Ltd.
- Hom, Christopher. 2010. "Pejoratives." *Philosophy Compass* 5 (2): 164–85. <https://doi.org/10.1111/j.1747-9991.2009.00274.x>
- Hom, Christopher. 2012. "A Puzzle About Pejoratives." *Philosophical Studies* 159 (3): 383–405. <https://doi.org/10.1007/s11098-011-9749-7>
- Hom, Christopher, and Robert May. 2013. "Moral and Semantic Innocence." *Analytic Philosophy* 54 (3): 293–313. <https://doi.org/10.1111/phib.12020>
- Hughes, Geoffrey. 2006. *An Encyclopedia of Swearing: The Social History of Oaths, Profanity, Foul Language, and Ethnic Slurs in the English-Speaking World*. New York: M. E. Sharpe.
- Jeshion, Robin. 2013. "Expressivism and the Offensiveness of Slurs." *Philosophical Perspectives* 27 (1): 231–59. <https://doi.org/10.1111/phpe.12027>
- Jeshion, Robin. 2016. "Slur Creation, Bigotry Formation: The Power of Expressivism." *Phenomenology and Mind* (11): 130–9. https://doi.org/10.13128/Phe_Mi-20113
- Nunberg, Geoff. 2018. "The Social Life of Slurs". In *New Work on Speech Acts* edited by Daniel Fogal, Daniel Harris, and Matt Moss, 237–95. Oxford: Oxford University Press. <http://doi.org/10.1093/oso/9780198738831.001.0001>
- Rappaport, Jesse. 2019. "Communicating with Slurs." *The Philosophical Quarterly* 69 (277): 795–816. <https://doi.org/10.1093/pq/pqz022>
- Schlenker, Philippe. 2007. "Expressive Presuppositions." *Theoretical Linguistics* 33 (2): 237–45. <https://doi.org/10.1515/TL.2007.017>
- Vallée, Richard. 2014. "Slurring and Common Knowledge of Ordinary Language." *Journal of Pragmatics* (61): 78–90. <http://dx.doi.org/10.1016/j.pragma.2013.11.013>

-
- Whiting, Daniel. 2013. "It's Not What You Said, It's the Way You Said It: Slurs and Conventional Implicatures." *Analytic Philosophy* 54 (3): 364–77. <https://onlinelibrary.wiley.com/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, 137–58. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195367881.001.0001>