

Slurs, Neutral Counterparts, and What You Could Have Said

Arianna Falbo

arianna_falbo@brown.edu

Forthcoming in *Analytic Philosophy*.

Penultimate draft.

Abstract: Recent pragmatic accounts of slurs argue that the offensiveness of slurs is generated by a speaker's free choice to use a slur opposed to a more appropriate and semantically equivalent neutral counterpart. I argue that the theoretical role of neutral counterparts on such views is overstated. I consider two recent pragmatic analyses, Bolinger (2017) and Nunberg (2018), which rely heavily upon the *optionality* of slurs, namely, that a speaker exercises a deliberate lexical choice to use a slur when they could have easily used a neutral counterpart instead. Against such views, I argue that across a range of different offensive uses of slurs, a speaker's choice to use a slur opposed to a neutral counterpart plays little to no role in accounting for why the slur generates offence. Such cases cast serious doubt upon the explanatory depth of these pragmatic analyses, and raise more general concerns for views which draw upon the relationship between a slur and its neutral counterpart. The main upshot is this: theorists should exercise caution in assuming that neutral counterparts play any fundamental or systemic role in explaining why slurs are offensive.

“If you can’t say something nice, don’t say nothing at all.”

— Thumper, Disney’s *Bambi*

1. Introduction¹

It has become commonplace in the literature on slurs to assume that the meaning of a slur is, to some degree or other, theoretically tethered to its neutral counterpart. This is especially true for pragmatic analyses of slurs, which argue that a slur is semantically equivalent to its neutral counterpart. For example, according to such views, the N-word is semantically equivalent with

¹ Content warning: this paper mentions a number of slurs throughout. I recognize that even mentions of slurs in academic philosophy papers, which discuss the offensiveness of slurs, can reasonably make some people uncomfortable, and may generate offence. I have tried my best to limit the mentions of slurs to only what is argumentatively and dialectically required.

and, hence, makes the exact same truth-conditional contribution as, its neutral counterpart: ‘African American.’

In what follows, I argue that the theoretical role of neutral counterparts in prevailing pragmatic analyses of slurs is overstated. As a result, such views are unable to explain why slurs are offensive across a range of cases. In section two, I give a brief overview of neutral counterparts and their role in prevailing accounts of the offensiveness of slurs. In section three, I survey the current literature and raise some initial worries for accounts which appeal to neutral counterparts in explaining the offensiveness of slurs. In section four, I focus on recent pragmatic analyses defended by Bolinger (2017) and Nunberg (2018). Both accounts rely heavily upon the *optionality* of slurs, namely, the idea that in choosing to utter a slur a speaker exercises a lexical choice to use a slur *opposed to* an unmarked and semantically equivalent neutral counterpart. In section five, I outline a series of offensive uses of slurs which are not adequately explained by these accounts. Across a range of different usages, a speaker’s choice to use a slur opposed to a neutral counterpart plays little to no role in accounting for why the slur is offensive. Such cases cast serious doubt upon the explanatory depth of the pragmatic approach in particular, and raise more general concerns for any analysis of the offensiveness of slurs which draws upon the relationship between a slur and its neutral counterpart.

The main upshot is this: theorists should exercise caution in assuming that neutral counterparts play any fundamental or systematic role in explaining why slurs are offensive.

2. Neutral Counterparts

It is widely assumed that slurs have neutral counterparts. Neutral counterparts are co-referential terms which refer to the same target group non-pejoratively. For example, Hornsby (2001) argues that:

...for each such word [i.e. each slur], there is, or at least perfectly well could be, another that applies to the same people but whose use does not convey these things—there is, that is, a neutral counterpart.²

Similarly, Vallée (2014) assumes the following principle:

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.³

Following suit, Camp (2013) suggests that a slur’s contrast with its neutral counterpart is a part of the very nature of what it is to be a slur. She says:

² Hornsby (2001): 129.

³ Vallée (2014): 79.

Indeed, a slur's *very optionality* is part of what makes it so expressively powerful—slurs are arguably *constituted as slurs* in part through their contrast with (comparatively) neutral counterparts.⁴

Also, Camp (2018):

...part of what is infuriating about a slur is its *optionality*: the speaker has thrown a volatile assumption into the conversation when they could just as easily have stuck with a neutral counterpart.⁵

The assumption that slurs have corresponding neutral counterparts is widespread in the literature and is endorsed among theorists who otherwise defend very different views. As such, the relationship between neutral counterparts and slurs often plays some role in explaining the offensiveness of slurs.⁶

Broadly three main kinds of accounts have emerged in recent scholarship. *Pragmatic* accounts defend a semantic equivalence thesis, namely, they argue that the meaning of a slur is identical to the meaning of its neutral counterpart. Given this synonymy, pragmatic theorists must appeal to extra-semantic mechanisms in order to account for the offensiveness of slurs.⁷

Semantic accounts standardly claim that slurs supplement derogatory content about the referenced group to the meaning of neutral counterparts. Accordingly, slurs are not synonymous with their neutral counterparts, but the meaning of a neutral counterpart is normally taken to be subsumed under the meaning of a slur. For example, the meaning of 'w*p' will contain the same semantic content as 'Italian,' but the former will also encode negative stereotypes or other discriminatory content concerning Italians.⁸

Lastly, *expressivist* accounts claim that slurs express contempt towards their targets because they are presumed to be a member of a certain social group. What differentiates slurs from their neutral counterparts is that slurs have an attitudinal or expressive component that their neutral correlates lack.⁹

Differences aside, nearly all accounts appeal, to some degree or another, to neutral counterparts in explaining the offensiveness of slurs. However, it is important to highlight that, when compared to other views, the role of neutral counterparts is far more pronounced on pragmatic analyses.

⁴ Camp (2013): 340, italics added.

⁵ Camp (2018): 41-2, italics added.

⁶ For further discussion see Croom (2011): 345-355.

⁷ See for example, Bolinger (2017), and Nunberg (2018) (both of these accounts will be discussed in detail in Section 4). Also, see Anderson and Lepore (2013a, 2013b).

⁸ See for example, Tirrell (1999), Hom (2008, 2010), Hornsby (2001), Williamson (2009), and Camp (2013, 2018).

⁹ See for example, Jeshion (2013a, 2013b, 2016, 2017), Potts (2007), Richard (2008), and Saka (2007). Also, cf. Hom (2010): 171 who describes expressivism about slurs as follows:

Where D is a derogatory word and N is its neutral counterpart, someone who predicates D of x, (i) says that x is N, and (ii) condemns those who are N.

This is because pragmatic accounts defend semantic equivalence between slurs and their neutral counterparts. The relationship between slurs and neutral counterparts need not be as close-knit on semantic and expressivist accounts. Because semantic accounts lexically encode distinct semantic content of some kind, it is open to defenders of such views to claim that the semantic meaning of a slur is independent and distinct from the meaning of its neutral counterpart.¹⁰

An analogous move is available to the expressivist. It is open to the expressivist to deny that slurs are semantically equivalent to their neutral counterparts, while nonetheless maintaining that the offensiveness of slurs flows from some attitudinal or expressive dimension of slurs.¹¹ Semantic and expressivist accounts are thus more easily able to jettison the assumption that slurs have a close theoretical relationship to neutral counterparts. This is primarily due to the fact that such views are not wedded to the assumption that slurs are semantically equivalent to neutral counterparts. So, among the three camps, pragmatic views seem to have the greatest reliance upon neutral counterparts in accounting for the offensiveness of slurs.

In the next section, I survey recent literature and raise some preliminary concerns for accounts which appeal to neutral counterparts when explaining the offensiveness of slurs.

3. Some Initial Worries

On the face of it, the claim that slurs are associated with neutral counterparts may seem innocuous or even obviously correct. However, upon closer scrutiny worries arise.

3.1 Do All Slurs Even Have Neutral Counterparts?

We might reasonably wonder whether all slurs even have corresponding neutral counterparts. Jeshion (2016) notes that slurs such as ‘hillbilly,’ ‘redneck,’ and ‘white trash’ fail to have neutral correlates. These terms are used primarily to refer to low-income and uneducated people living in rural areas of the Southern United States (primarily in the Ozarks and Appalachia).¹² ‘Southerner’ is not a plausible neutral counterpart because it is too generic and fails to be co-referential.

Jeshion also draws our attention to a class of slurring terms which are used to refer wholesale to individuals from disparate social groups. For example, ‘d*go’ is used to refer to a person of Italian, Portuguese, Spanish, or other Mediterranean descent.¹³ In such cases, there is a many-to-one relation such that for one slurring term there is multiple associated neutral counterparts. This data is incongruous with a pragmatic view which assumes semantic equivalence between slurs and neutral counterparts.

¹⁰ For a view which deploys this strategy see, for example, Neufeld (2019).

¹¹ For views which exhibit this approach see, for example, Saka (2007): 140-143, and Richard (2008). For an expressivist view which assumes truth-conditional equivalence between slurs and neutral counterparts see Jeshion (2013a, 2017).

¹² Jeshion (2016): 135.

¹³ Jeshion (2016): 135 also notes that “w*g” is used to refer to “anyone of Middle Eastern, south Asian, Eastern European, and Mediterranean descent.”

DiFranco (2015) has also argued that metaphorical or figurative slurs, such as ‘Jewish American Princess’, fail to have neutral counterparts. ‘Jewish American Woman’ is too broad and fails to co-refer. This is because it does not capture that this is a slur used by speakers to convey that its targets are high maintenance, spoiled, pampered, and so on.¹⁴ Additionally, slurs targeted at the LGBTQ+ community such as ‘fairy’ and ‘twink’ face similar over-simplicity problems and failures of co-reference when correlated with the neutral correlate ‘gay.’

3.2 Neutrality?

Focusing on gendered slurs, Ashwell (2016) argues that slurs like ‘sl*t’ lack properly *neutral* counterparts. ‘Woman’ is insufficient seeing as ‘sl*t’ is used primarily toward women in virtue of some presumed disposition towards sexual promiscuity. Moreover, Ashwell argues that the description ‘woman who is promiscuous’ (or some variant thereof) has *normative* content and is thus not sufficiently neutral. Promiscuity is not socially acceptable, so calling a woman “promiscuous” assumes that there is some set number of sexual partners that she ought to have.¹⁵ Hence, Ashwell argues that this description lacks the neutrality needed to count as a genuinely *neutral* counterpart.

One might object, contra Ashwell, and argue that while some slurs may not have correlates which are free from normative evaluation, they do typically (or even always) have non-offensive correlates or at least correlates which don’t offend in the same way or to the same degree that slurs do. For example, one might argue that ‘woman who is promiscuous’ is a neutral correlate of ‘sl*t’ because even if the former generates offence, it’s an offence of a lesser kind which is sufficiently different than the offence generated by the associated slur. So, Ashwell is right that expressions like ‘woman who is promiscuous’ do plausibly express normative judgments concerning the number of sexual partners a woman ought to have, but this is compatible with its being a (relatively) neutral counterpart. According to this line of response, neutral counterparts need not be entirely free of normative elements, they just can’t convey the same kind and degree of offence that slurs do.¹⁶

This reply to Ashwell highlights a general need in the literature to clarify what “neutral” means when theorists claim that slurs have “neutral counterparts.” Neutral in what sense? Is a neutral counterpart properly neutral if it nonetheless incorporates normative elements? In general, those who make use of neutral counterparts in their theories of slurs typically don’t offer a precise definition of neutrality.¹⁷

These are undoubtedly delicate and complex issues, and I don’t intend to defend a specific definition of neutrality here either. However, nothing in what follows turns on this. For the sake

¹⁴ See DiFranco (2015): 32 for various other examples of these kinds of slurs. DiFranco argues that once we move away from considering single-term noun phrases, and shift to considering the phenomena of slurring more widely, that the presumed connection between slurs and neutral counterparts is rendered less plausible.

¹⁵ Ashwell (2016): 235.

¹⁶ Thank you to an anonymous referee for raising this objection.

¹⁷ See Nunberg (2018): 262-274 for a related discussion.

of argument, we can put definitions of neutrality aside and proceed with just the following necessary condition: a neutral counterpart must fail to convey the same kind of offence or negative evaluation of its referent as the corresponding slur does.

3.3 Semantic Equivalence

Pragmatic accounts are committed to the thesis that a slur is semantically equivalent to its neutral counterpart. Accordingly, statement like (1) are interpreted as conceptual truths, akin to statements like ‘Every vixen is a female fox.’

(1) All Jews are k*kes.

Many view this as a bad result. Croom (2015) and Nufeld (2019) have also argued that the semantic equivalence thesis is unable to account for data such as the following, what we might call *non-target* uses of slurs:

(2) Deandra’s a lesbian, but she’s not a dy*e.

(3) I love Black people, but I hate n*****.¹⁸

Given the semantic equivalence thesis, these utterances should sound incoherent and contradictory. But (2) and (3) don’t appear to be contradictions. They are clearly distinct from utterances such as the following.

(4) Thomas is a vegan, but he eats bacon.

(5) My neighbour is a bachelor, but he’s married.

Non-target uses of slurs aren’t fringe or anomalous, and they ought to be accounted for by any analysis of the offensiveness of slurs.

In a similar vein, we can consider the following data, what we may call *reinforcing* uses of slurs.

(6) Not only is my neighbour a lesbian, but she’s a total dy*e too. She drives a motorcycle and is covered in tattoos.

(7) Of course my soccer coach is Italian, and he’s a total w*p too. He eats a prosciutto sandwich during every practice, and I wouldn’t be surprised if he was in the mob.

Or consider the following heartbreaking example. Imagine a father saying the following to his son.

(8) I don’t care that you’re gay, you will always be my boy. I just wish you weren’t such a f***ot.

Data like (6)-(8) are analogous to non-target uses.¹⁹ However, according to semantic equivalence, just as non-target uses like (2) and (3) are interpreted as contradictory, on the other side of the coin, utterances like (6)-(8) are deemed trivial or redundant. But they don’t appear to be. Thus, this data casts serious doubt upon the thesis that a slur is semantically equivalent to its neutral counterpart.²⁰

¹⁸ This is from Chris Rock’s 1996 HBO comedy special *Bring the Pain*.

¹⁹ Also compare Louis C.K.’s (2008) Comedy Special, *Cheved Up*: “I would never call a gay guy a ‘f***ot’, unless he was being a ‘f***ot, but not because he’s gay.” For further discussion of these sort of cases also see Jeshion’s (2013a) discussion of “G-extending” and “G-contracting” uses of slurs.

²⁰ There is yet further data against the semantic equivalence thesis. For example, if a slur is equivalent to its neutral counterpart, then we should be able to substitute it, *salva veritate*, but in many cases this does not seem feasible. This worry is especially salient in the case of belief reports. Consider the following two statements:

Accounts that posit neutral counterparts thus face notable hurdles in capturing a range of slurring phenomena. This is because some slurs may fail to have corresponding neutral counterparts altogether or may fail to have a single unique and, hence, truth-conditionally equivalent, neutral counterpart (as in the case of slurs which refer wholesale to multiple target groups). There are also general explanatory challenges concerning the relevant notion of neutrality that is at issue in this debate and what it even means to be a properly *neutral* counterpart. Furthermore, a range of data speaks against the semantic equivalence thesis.

I survey these issues solely in order to raise some initial doubts concerning the overall prospects of appealing to neutral counterparts when explaining why slurs are offensive. The arguments against the pragmatic view which follow, however, do not rest upon the claim that slurs lack neutral counterparts, or that neutral counterparts aren't appropriately neutral, or that slurs are not semantically equivalent to their corresponding neutral counterparts. So, for the time being, let's grant that (i) all slurs do in fact have neutral counterparts, (ii) that such neutral counterparts are neutral insofar as they (at minimum) don't convey the same kind and degree of offense as the corresponding slurs do, and (iii) that slurs and their neutral counterparts are semantically equivalent.

In the next section, I outline two recent pragmatic analyses of the offensiveness of slurs defended by Bolinger (2017) and Nunberg (2018).

4. Pragmatic Accounts

Bolinger (2017) and Nunberg (2018), describe a wholly pragmatic mechanism which, they argue, explains the offensiveness of slurs. At the heart of both of their accounts is the *optionality* of slurs, namely, the idea that when using a slur a speaker exercises a lexical choice in that they could just have easily used a neutral counterpart expression instead. I explain each of these views in turn.

4.1 Contrastive Choice

Bolinger (2017) argues that the offensiveness of slurs consists in the fact that the speaker decided to use a slurring term opposed to a neutral counterpart.

In choosing to use a slurring term *rather* than its neutral counterpart, the speaker signals that she endorses the term (and its associations). Such an endorsement warrants offense, and consequently slurs generate offense whenever a speaker's use demonstrates a contrastive preference for the slurring term.²¹

(A) 'I believe that Frank is Jewish.'

(B) 'I believe that Frank is a k*ke.'

Intuitively, the truth value of (A) and (B) can come apart. This, however, is incompatible with the assumption that slurs and their neutral counterparts are synonymous. For further discussion see, for example, Croom (2015), Hom (2009, 2010) and Neufeld (2019).

²¹ Bolinger (2017): 439.

For example, by choosing to use the expression ‘w*p’ instead of ‘Italian’ a speaker signals to her interlocutor that she endorses the use of this term and (at least some of) the negative stereotypes and associations connected to it. In using a slurring term, a speaker freely chooses to utter a blatantly marked expression, opposed to an unmarked salient alternative. More schematically, Bolinger says:

For some content φ , when it is common knowledge in the linguistic community that:

(i) a is an expression for ψ associated with φ , and

(ii) β is an expression for ψ not associated with φ ,

then in situations where the choice of expression is not forced, and the speaker is aware of (i) and (ii), selecting a in contrast to β signals that the speaker endorses or shares φ .²²

According to Bolinger slurs are offensive because the speaker had the choice between two expressions, (a or β), one of which is associated with bigotry (the slur) and one which is not (the neutral counterpart). The fact that the speaker has freely chosen to use the slur sends a signal to her interlocutor that she endorses the associated bigotry (φ). Bolinger notes that this signal can be neutralized if it becomes common knowledge that the speaker is genuinely unaware that a is associated with φ , or if the speaker didn’t know that she could have used β instead which was not associated with φ .²³ Additionally, if the use of a is forced such that the alternative expression β is unable to be used, then the speaker’s use of a does not signal φ . An example of a forced context might be the use of a slur in direct speech report, for instance, when a speaker is offering witness testimony in a court trial and recounting the precise and exact words may really matter.²⁴ Also, it is important to note that in the above formulation Bolinger assumes that a and β are co-referential and synonymous: they are both expressions for ψ . Hence, according to this account slurs are synonymous with, and hence, pick out the same extension as their neutral counterparts.

To summarize, according to the *Contrastive Choice Account*, the offence generated by slurs is accounted for by a purely pragmatic mechanism: it consists in a signal that is sent to one’s interlocutor as the result of the speaker’s choice to use a slurring term *in contrast to* or *opposed to* an unmarked and co-referential neutral counterpart. This signal can be cancelled in certain contexts,

²² Bolinger (2017): 447.

²³ Cf. Bolinger (2017): 448 “when a speaker is ignorant of the association between a and β , her use of a does not signal that she endorses φ . If she is unaware of any alternatives to a , then whether her use warrants offense depends on whether, if she had known of an alternative β , she would have refrained from using a .”

²⁴ Though, Bolinger (2017): 451-2 and fn. 11, notes that even here offense might still be warranted if the speaker could have used a circumlocution, e.g., “N-word” or “C-word”. Such offence is evident in the recent firing of a Netflix communications executive who used a slurring term in two meetings while discussing the use of sensitive language, see Wattles (2018). Another example is the cancellation of a Princeton University class (an anthropology course entitled: “Cultural Freedoms: Hate Speech, Blasphemy and Pornography”) following the professor’s use of the N-word in full, see Flaherty (2018a, 2018b). Given that many slurs fail to have salient and unambiguous circumlocutions of this form, I do not consider these alternative options throughout the paper, though they certainly warrant further investigation.

namely, when it is common knowledge that the speaker is ignorant of the fact that *a* is associated with φ or when the speaker is not free to use β instead of *a*.²⁵

4.2 Ventriloquist Implicature

Nunberg (2018) argues that the offensiveness of slurs is the result of a conversational implicature, specifically, what he calls a *ventriloquist implicature*. Such implicatures arise when a speaker blatantly deviates from the default conventions of one's conversational context, and instead chooses to use a marked referring expression which has bigoted affiliations. As a result, the speaker violates the Maxim of Manner which, in Grice's (1989) words, relates to: "*how* what is said is to be said."²⁶ Grice's sub-maxims of Manner include the following: avoid ambiguity, avoid obscurity, be brief or succinct, and be orderly. To this list, Nunberg adds the sub-maxim: *be appropriate*. Following this sub-maxim requires a speaker to use language in accordance with the appropriate default lexical conventions of one's conversational context. Ventriloquist implicatures arise:

...only when the word used by the speaker replaces one prescribed by the convention that would be the contextual default. In other words, there has to be a normal conventional (i.e., a lexicalized) means of saying what the speaker is saying abnormally.²⁷

When a speaker decides to use a slurring term opposed to a salient nonpejorative and conventionalized default expression, that is, instead of a neutral counterpart, they transgress against the default linguistic conventions of the conversation. The speaker outwardly indicates a preference for a pejoratively marked alternative expression. And, as a result, the speaker thereby signals an affiliation with the users of the slur and its provenance. Nunberg describes this violation of the sub-maxim *use appropriate language* as follows:

...[the sub-maxim] implies an obligation to *make appropriate choices* among the welter of conventions that might govern the choice of words in a given speech situation, depending on the social norms that are contextually pertinent. Opting out of these maxims can set up various conversational implicatures...redskin is distinguished from Indian not by any additional evaluative or expressive features of its meaning, but merely in being the description of Indians prescribed by the conventions of a group whose members have disparaging attitudes about American Indians. Then the implications of pointedly choosing to use redskin arise not from the meaning of the word but from its association with the discourse of a certain group of speakers.²⁸

²⁵ Additionally, there are appropriated or reclaimed uses of slurs wherein a slur does not signal offence. See Bolinger (2015): 455-56 for further discussion. Consideration of reclaimed uses of slurs is beyond the scope of the paper. For discussions of appropriated uses of slurs see, for example, Anderson (2018).

²⁶ Grice (1989): 27.

²⁷ Nunberg (2018): 268.

²⁸ Ibid: 244.

On Nunberg's account, slurs have a self-affiliating function: by using a slur a speaker signals "his affiliation with the group and its point of view."²⁹ For example, by using the term 'w*p' opposed to the neutral counterpart 'Italian', a speaker self-identifies or self-affiliates with the bigoted attitudes of a community of language users who own and use this term to refer to Italian people. Just as a ventriloquist creates the illusions that a dummy's voice is coming from elsewhere, by using a slur a speaker imitates and invokes the lexical conventions a bigoted community. In a slogan: bigots don't use slurs because they're offensive; slurs are offensive *because* they're the words that bigots use.³⁰

To summarize, according to Nunberg's *Ventriloquistic Implicature Account*, the offence generated by slurs happens at the level of a pragmatic inference: it is the result of a conversational implicature or as Nunberg calls it: a ventriloquist implicature. The offensiveness of a slur is contingent upon the existence of a co-referential term that does not convey the same negative evaluation of its referent. The potential for offense arises when a speaker is confronted with a lexical choice between two semantically equivalent expressions. On the one hand, they can use the appropriate conventionalized default term which is customarily used to refer to the target group (e.g., 'Italian'). Or, on the other hand, they can opt to use an unorthodox marked alternative referring expression (e.g., 'w*p'). By freely choosing to use the marked alternative, a speaker signals an affiliation with a bigoted community and, as a result, the utterance generates offence.

It is important to highlight the indispensable role neutral counterparts have on these analyses. On both accounts, the offensiveness of slurs consists, to a large extent, in the fact that a speaker freely chooses to employ a slurring term *opposed to a neutral counterpart*. The slurring speaker transgresses against the default lexical conventions that are operative in the conversation. This transgression pragmatically implicates that the speaker endorses the derogation of the referenced group³¹ and that they self-affiliate with an associated bigoted community who owns and uses the term.³²

There are important differences between these two accounts, but where these views overlap is what's pertinent to our purposes. In what follows, I treat both of their accounts uniformly, and for brevity I will refer to them jointly as *the pragmatic account*. This will undoubtedly obscure and oversimplify important nuances and differences between these views, but this loss does not impinge upon our main goal, namely, to explore the theoretical role and limits of neutral counterparts in prevailing pragmatic explanations of the offensiveness of slurs.

In the next section, I discuss a range of offensive uses of slurs which are not adequately explained by the pragmatic account. This is because in such cases the speaker is not plausibly interpreted as choosing to use a slurring term *opposed to a salient neutral counterpart*.

²⁹ Ibid: 38.

³⁰ Ibid: 244.

³¹ Bolinger (2017): 448.

³² Nunberg (2018): 261.

5. What You Could Have Said

On the pragmatic account, when a speaker chooses to employ a slurring term, she explicitly opts for an expression that is markedly derogatory, opposed to a salient unmarked and co-referential alternative: a neutral counterpart. Nunberg argues that slurs are marked in virtue of being the words that bigoted communities use to refer to the targeted group. Bolinger argues that a slur is marked in virtue of being used primarily by speakers who endorse the oppressive treatment of certain groups.³³ Speakers are confronted with a pointed lexical choice—they can stick with orthodox conventions and use a conversationally appropriate neutral counterpart, or they can transgress and use a slur. Transgressions generate offence.

To understand the conversational implicature responsible for the offensiveness of slurs on the pragmatic account, it is helpful to first consider Manner implicatures in non-slurring contexts. Imagine a parent, Artemis, who utters the following to her partner while in the presence of their young child.

(9) Should I pick up some I-C-E C-R-E-A-M from the store?

In uttering (9) Artemis implicates that she does not want their child to hear the expression ‘ice cream.’ Artemis masks her message by opting to using a more obscure expression, ‘I-C-E C-R-E-A-M.’ This results in a flagrant violation of Grice’s Maxim of Manner, specifically, the sub-maxims: avoid obscurity and be brief or succinct.

What pragmatic reasoning leads Artemis’ interlocutor to the conclusion that the speaker does not want the child to hear the word ‘ice cream’? According to Grice, communication is primarily a cooperative undertaking which is governed by a range of expectations or maxims concerning how conversation should be carried out. Interlocutors are expected to offer contributions to the conversation which are informative, truthful, relevant, and appropriate.³⁴ Hence, in light of this, we can imagine Artemis’ partner reasoning as follows in response to (9):

(i) Instead of uttering (9), Artemis could have said:

(10) Should I pick up some ice cream from the store?

(10) would have been more cooperative because it’s truth-conditionally equivalent to (9), while being less obscure and more succinct. But Artemis did not utter (10), so why did she utter (9) instead?

(ii) The most plausible explanation for Artemis’ uttering (9), opposed to (10), is that (10) was, for some reason, inappropriate or unavailable to utter in the context.

(iii) The most plausible reason for why Artemis couldn’t utter (9) instead of (10) is because she didn’t want their child to hear the word ‘ice cream.’

The precise details of the pragmatic reasoning in this case are mostly orthogonal to our purposes. What is relevant is the very first stage of reasoning concerning what Artemis *could have said*. The option set that Artemis draws from can be understood in terms of a set of referring

³³ Note that Bolinger (2017) also mentions in fn. 32 that this is not the *only* way a slur can be marked, but nonetheless claims that this is the primary way.

³⁴ Grice (1989): 26-7.

expressions that are in pragmatic competition with each other. In this particular case, the set of salient alternative expressions plausibly includes: <I-C-E C-R-E-A-M, ice cream>. Under the assumption that one's interlocutor is observing Gricean Maxims, one can plausibly reason from the fact that a speaker used a more obscure referring expression ('I-C-E C-R-E-A-M') *opposed to* a less obscure and truth-conditionally equivalent expression ('ice cream'), to the conclusion that the speaker wanted to mask her message so that the child did not hear it.

Calculating conversational implicatures, and specifically Manner implicatures, requires one to interpret the speaker as exercising a choice between salient candidate expressions. One must recognize that a speaker has chosen to employ a marked expression opposed to an unmarked alternative which could have plausibly been used instead. Recognizing the contrastive choice that the speaker faces is precisely what generates the conversational implicature.

Let's return to the pragmatic account of the offensiveness of slurs. Recall that according to this account what makes a slur offensive is that the speaker could have used a neutral counterpart instead of slurring term. Accordingly, when an interlocutor is in the process of computing a pragmatic implicature and is considering why the speaker decided to use a slurring term, the set of competing expressions always has the following form: <NEUTRAL COUNTERPART, SLUR>. The hearer recognizes that the speaker could have used a term that was equally as informative and significantly more appropriate given the lexical conventions operative in one's conversational context, but decided not to. This implicates that the speaker has a preference for a term which derogates its target and has bigoted affiliations.

The pragmatic account gives theoretical priority to neutral counterparts because it assumes that the choice a speaker is confronted with is *always* between a slurring term and its neutral counterpart. This assumption, however, fails to properly capture the lexical choices of speakers and the attendant pragmatic reasoning of their interlocutors across a diverse range of cases.

5.1 Demonstrative Uses

Consider the following, what I will call *demonstrative uses* of slurs. The communicative situation is as follows: Charlie and Dennis, two friends, are discussing the end of Obama's term as President at the local pub.

(11) Charlie: What is he doing now?

(12) Dennis: That n***** is on vacation.

How should Charlie interpret Dennis' use of the N-word? Or more precisely, how might we calculate the conversational implicature that generates offence in this case according to the pragmatic view? What alternative expressions could have been uttered instead of the slur? The pragmatic analysis suggests that we interpret Charlie's pragmatic reasoning as follows:

(i) Instead of uttering (12), Dennis could have said:

(13) That African American is on vacation.

But Dennis did not say (13), which would have been more appropriate given the default lexical conventions of their conversation, so why did he say (12)?

(ii) The most plausible explanation for why Dennis uttered (12) is that he wanted to derogate African Americans and affiliate himself with a community of racists who use the N-word refer to African Americans.

But if Dennis had replied with (13) this would have been very awkward and incredibly unnatural. Additionally, uttering (13) instead of (12) is likely to preserve offense, not neutralize it. However, this offense is likely of a different kind and lesser severity than the offense generated by the slur; it is the sort of offense that may arise from needlessly mentioning demographic information.³⁵

It is far better to interpret the first phase of reasoning, where Charlie is considering what Dennis could have said instead, as not including ‘African American’ among the salient alternative expressions at all. More plausibly, Dennis could have simply said ‘he’ or ‘Obama’ instead of using a slur. These expressions are the most obvious unmarked alternatives in this context. The use of a relevant pronoun or name would have made Dennis’ utterance far less linguistically stilted, compared to (13), and wouldn’t have resulted in a transgression or explicit flouting of the Maxim of Manner.

In demonstrative uses of slurs, then, the relevant set of salient alternative expression is often not of the form <SLUR, NEUTRAL COUNTERPART>, but is more plausibly: <SLUR, PRONOUN/NAME>. What generates offence in these cases is not that the speaker decided to use a slur when they could have easily used a neutral counterpart instead, but that they chose to use a slur instead of a more conversationally appropriate name or pronoun.

One might raise the following response on behalf of the pragmatic view. One may agree that in demonstrative uses of slurs, pronouns and names are typically *the most* salient and unmarked alternative expressions to use instead of the slur. However, this doesn’t entail that neutral counterparts are wholly unavailable to speakers as alternative expressions—neutral counterparts still are conversationally salient alternatives, they are just not *as salient* as pronouns and names. Linguistic salience, one might claim, isn’t as coarse-grained as I appear to have assumed. Instead, linguistic salience might be better understood on a continuum and as coming in degrees. So, what’s crucial to the pragmatic account, and specifically to the explanation of why demonstrative uses of slurs are offensive, is only that the neutral counterpart term be closer to optimal salience relative to the slur. Hence, understood in this way, neutral counterparts may still be used to explain the offensiveness of slurs, even if there are other alternative expressions (e.g., names, pronouns) which might be more salient in the context.³⁶

It is certainly true in general that there are often a range of salient alternative expressions that a speaker could have uttered in a specific context, some of which are salient to a greater degree

³⁵ To motivate this idea it is perhaps helpful to consider the following conversational exchange:

(A): Who’s coming to dinner?

(B): Thomas and he’s bringing his Black girlfriend too.

Mentioning that Thomas’ girlfriend is Black suggest that this aspect of her is somehow relevant to the conversation. When this demographic information is needlessly mentioned, it can often generate offense.

³⁶ Thank you to an anonymous referee for raising this objection.

than others. However, even if we grant the assumption that neutral counterparts are among the set of salient alternative expressions in demonstrative uses of slurs, worries still arise for the pragmatic view. In any case, the pragmatic theorist must give a plausible account of an algorithm or method of reasoning whereby the listener can work out the relevant implicature.³⁷ Specifically, they must explain how a listener may reasonably be led to the view that the speaker could have used a neutral counterpart instead of a slurring term. But it's unclear, according to the pragmatic view, what this reasoning might be.

I will consider two responses on behalf of the pragmatic view for how a listener might pragmatically reason to the claim that the speaker should have used a neutral counterpart instead of the slur. I argue that the first option is ultimately untenable and that the second option, while more plausible, is incompatible with a pragmatic approach.

Here's the first option. When considering what the speaker could have said instead of the slur the listener might appeal to an open-ended or fairly wide range of salient alternative expressions, which includes the neutral counterpart as well as the relevant name and pronoun.³⁸ However, reasoning from this broad starting point is unlikely to sufficiently zero in on the neutral counterpart expression. This is simply for the reason that the relevant pronoun and name are less marked and, hence are more appropriate and salient alternatives when compared to the neutral counterpart. Thus, interpreting the computation of the conversational implicature in this way is unlikely to lead the listener to the view that the speaker could have used a neutral counterpart instead.

Here's another option for how the listener might reason to the view that the speaker could have reasonably used a neutral counterpart instead of the slur. The listener's reasoning might be lexically specified with a conventionalized scale.³⁹ To motivate this option, consider the following example of a standard scalar implicature. Imagine a speaker who utters the following.

(14) Mac ate some of the pie.

(14) implicates that Mac didn't eat all of the pie. The standard pragmatic reasoning which leads the hearer to this conclusion can be glossed as follows.

(i) Instead of uttering (14), the speaker could have said (15), which would have been more informative.

(15) Mac ate all of the pie.

But the speaker did not say (15), so why did they say (14)?

(ii) The most plausible explanation is that the speaker was not in a position to utter (15).

And the most likely reason for this is because they don't believe it. So, Mac mustn't have eaten all of the pie.

We can interpret the interlocutor in this case as drawing upon the following conventionalized scale: <some, all>. It is important to note that scalar implicatures rely upon the semantic meanings of the terms used; in this case the meaning of 'some' and the meaning of 'all.'

³⁷ Cf. Grice (1989): 50.

³⁸ This option reflects a broadly Gricean (1989) approach to computing conversational implicatures.

³⁹ Horn (1984), Hirschberg (1985), and Levinson (2000).

Let's return to demonstrative uses of slurs. Perhaps we can interpret the listener as reasoning from a similar kind of conventionalized lexical scale, where each lexical item is ordered in terms of its pejorative strength. So when considering the utterance of (12), the listener's reasoning would be based upon the assumption that, given the meaning of the N-word and given the meaning of 'African American', it would have been more appropriate to use the less marked expression (i.e. the neutral counterpart) opposed to the pejoratively marked term (i.e. the slur). And because the speaker didn't opt to use the unmarked and non-pejorative expression, they must have wanted their utterance to convey offence.

But, recall that a key commitment of the pragmatic view is that slurs are synonymous with their neutral counterparts. However, analyzing the pragmatic calculation in these cases as involving a conventionalized scale would hard-wire in *semantic* difference between slurs and neutral counterparts. Hence, the pragmatic view can't explain the pragmatic computation as involving a conventionalized lexical scale. This would go against a core commitment of the view: the semantic equivalence thesis. This option is thus not compatible with the pragmatic account.

So, even if neutral counterparts are in some sense available to listeners, in virtue of being more salient and conversationally appropriate than slurs, it remains unclear how a listener might be led to the view that the speaker could have reasonably used a neutral counterpart expression instead. To be clear: I don't take this to be a decisive strike against the pragmatic view, rather it poses a serious explanatory challenge. It is thus up to the pragmatic theorist to fill in this explanatory lacuna. However, the prospects of doing so, while also preserving a key theoretical role for neutral counterparts, seems highly unlikely.

5.2 Weapon Uses

Let's consider a different kind of case, what are often called *weapon uses* of slurs. The communicative situation is as follows. A gay couple is walking, hand-in-hand, down the street when an unprovoked a stranger yells out the following to them:

(16) Hey f***ots!

Consider another recent and real-life example. In July 2020, United States Congresswoman Alexandria Ocasio-Cortez was accosted by Congressman Ted Yoho on the steps outside of the Capitol. Yoho stuck his finger in Ocasio-Cortez's face and (among other things) called her a "f**king b*tch."⁴⁰

According to the pragmatic account, weapon uses of slurs are offensive because the stranger and Yoho exercised a free choice between uttering a slurring term, opposed to a salient neutral counterpart. And, because they opted for the former over the latter, their utterance generated offence. However, if the stranger had yelled out "Hey gay men!" or if Yoho had said "you're a woman"—the respective neutral counterparts—these utterances would be incredibly strange and unnatural. What's more, they are likely to still communicate offense, just of a different kind and to

⁴⁰ Grey (2020).

a lesser degree. So, contrary to the pragmatic account, it can't be that weapon uses of slurs are offensive because the speaker could have easily used a neutral counterpart instead.

Failing to use a neutral counterpart does not explain why weapon uses of slurs are offensive. More plausibly, and far less awkwardly, the speaker should have simply kept their mouth shut, remained silent, and minded their own business. In the case of weapon uses, the most salient alternative to slurring is often *silence*. The relevant question to consider is not why the speaker decided to use a slur opposed to a neutral counterpart, or more generally, why they decided to utter a slur opposed to some other more conversationally appropriate expression. Rather, we should ask: why did the speaker choose to hurl a slur at fellow human being in the first place, when they could have just not? This suggests that in the case of weapon uses, the relevant set of salient alternatives does not concern a *lexical* choice between salient expressions, but is more plausibly construed as follows: <SLUR, \emptyset > where ' \emptyset ' represents silence.⁴¹ The decision that the speaker confronts in these cases is not best described as a choice between competing referring expressions at all, but instead a choice between competing courses of action: to derogate and assault, or to keep quiet and not engage in conversation.

Weapon uses are among the most toxic and offensive ways (if not *the* most offensive way) that slurs can be used to harm others. Accounting for data from weapon uses is thus central to any explanation of the offensiveness of slurs. But the pragmatic account doesn't even appear to be designed, in the first place, to accommodate them. This is because the pragmatic reasoning it identifies is overly narrow in its application. It assumes that when uttering an offensive slur the speaker is always best interpreted as making a pointed lexical choice between various salient referring expressions which could have been felicitously uttered in the context. However, weapon uses draw our attention to an earlier point in time; to the initial decision to confront and engage in conversation.

When considering why weapon uses of slurs are offensive the pertinent contrast is not with what the speaker could have *said* instead, but with *why* the speaker choose to engage, to confront, to accost, or to assault, when they could have just not. Hence, the offence generated in these cases is not a function of the speaker opting to use a slur when they could have easily used a more salient and conversationally appropriate neutral counterpart. Speakers who utter weapon uses of slurs simply shouldn't have sparked a confrontation in the first place—the most salient alternative is not even an expression at all, but rather choosing not to express one's self.

Given the structure of the pragmatic view—specifically, its description of speakers as making a pointed lexical choice between salient expressions—it appears unable to explain why weapon uses of slurs are offensive. Without a major revision to the overall analysis, the pragmatic account is unable to adequately explain why weapon uses are offensive. And, once revised, it is

⁴¹ This analysis of demonstrative uses resonates with the prohibitionist account defended by Anderson and Lepore (2013a, 2013b). It is not my intention to defend such a view here. I only wish to appeal to silence as the most salient and appropriate alternative to slurring, specifically in the case of weapon uses of slurs. For a criticism of the prohibitionist view see, for example, Popa-Wyatt and Wyatt (2017).

highly unlikely that neutral counterparts will retain any special role in accounting for why weapon uses of slurs are offensive.

More generally, what these cases reveal is that among the domain of referring expressions, a neutral counterpart is not always or even typically a salient unmarked alternative across a range of contexts. The pragmatic account thus fails to acknowledge a variety of unmarked salient alternatives beyond neutral counterparts. This is because the pragmatic analysis has its mechanistic foundations in the presumed *optionality* of slurs—the idea that in choosing to use a slur one exercises a free lexical choice to utter a slur opposed to a neutral counterpart. But, as we have seen, there are a diverse range of alternative unmarked options available to speakers. And, across a variety of usages, neutral counterparts fail to figure as the most salient unmarked alternative expression to using a slur, and often fail to be a salient option entirely. Thus, against the pragmatic account, neutral counterparts don't enjoy any theoretical priority or special role when it comes to explaining the offensiveness of slurs.

6. Conclusion

According to recent pragmatic views, the offensiveness of slurs is a function of a speaker's deliberate choice to use a slurring term opposed to a more conversationally appropriate and unmarked neutral counterpart. However, there are various offensive uses of slurs where the most salient unmarked alternative expression is decidedly *not* a neutral counterpart. In many cases, the speaker could have chosen to use a pronoun or a name instead of the slur, or could have simply decided to not engage in conversation the first place. While I have limited my discussion primarily to recent pragmatic accounts, I hope that my arguments may serve as a more general caution against appealing to neutral counterparts when explaining the offensiveness of slurs.⁴²

References

- Anderson, L. (2018). Calling, Addressing, and Appropriation. In D. Sosa (Ed.), *Bad Words* (pp. 7-28). Oxford: Oxford University Press.
- Anderson, L., & Lepore, E. (2013a). Slurring Words. *Noûs*, 47, 25-48.
- Anderson, L., & Lepore, E. (2013b). What Did You Call Me? Slurs as Prohibited Word. *Analytic Philosophy*, 54, 350–363.
- Ashwell, L. (2016). Gendered Slurs. *Social Theory and Practice*, 42, 228-239.
- Bolinger, R. (2017). The Pragmatics of Slurs. *Noûs*, 51, 439-462.
- Camp, E. (2013). Slurring Perspectives. *Analytic Philosophy*, 54, 330-349.
- Camp, E. (2018). A Dual Act Analysis of Slurs. In D. Sosa (Ed.), *Bad Words* (pp. 29-59). Oxford: Oxford University Press.
- Croom, A. (2011). Slurs. *Language Sciences* 33, 343-358.

⁴² I'm grateful to an anonymous reviewer from *Analytic Philosophy*, as well as Zach Barnett, Thomas Brandt, David Christensen, Matthew Guerrero, Haigen Messerian, and Anna Tsvetkov. Special thanks to Scott AnderBois.

- Croom, A. (2015). The semantics of slurs: A refutation of coreferentialism. *Ampersand: An International Journal of General and Applied Linguistics*, 2, 30-38.
- DiFranco, R. (2015). Do Racists Speak Truly? On the Truth-Conditional Content of Slurs. *Thought: A Journal of Philosophy*, 4, 28-37.
- Flaherty, C. (2018a). Ending a Course Over the N-Word. *Inside Higher Ed*, February 14, 2018. Retrieved from:
<https://www.insidehighered.com/news/2018/02/14/princeton-professor-whowas-criticized-using-n-word-class-hate-speech-cancels-course>
- Flaherty, C. (2018b). The N-Word in the Classroom. *Inside Higher Ed*, February 12, 2018. Retrieved from:
<https://www.insidehighered.com/news/2018/02/12/two-professors-differentcampuses-used-n-word-last-week-one-was-suspended-and-one>
- Grey, E. (2020). 'F**king Bitch' And The Everyday Terror Men Feel About Powerful Women. *HuffPost*, July 23, 2020.
- Grice, H.P. (1989). *Studies in the Way of Words*, Cambridge: Harvard University Press.
- Hirschberg, J. (1985). *A theory of scalar implicature*. University of Pennsylvania, PhD Thesis.
- Hom, C. (2008). The Semantics of Racial Epithets. *Journal of Philosophy*, 105, 416-440.
- Hom, C. (2010). Pejoratives. *Philosophy Compass*, 5, 164-185.
- Horn, L. (1984). Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature. In D. Schiffrin (Ed.), *Meaning, Form and Use in Context*. Washington, D.C.: Georgetown University Press.
- Hornsby, J. (2001). Meaning and Uselessness: How to Think about Derogatory Words. *Midwest Studies in Philosophy*, 25, 128-141.
- Neufeld, E. (2019). An Essentialist Theory of the Meaning of Slurs. *Philosophers' Imprint*, 19, 1-29.
- Nunberg, G. (2018). The Social Life of Slurs. In D. Fogal, D. Harris, & M. Moss (Eds.) *New Work on Speech Acts* (pp. 237-295). Oxford: Oxford University Press.
- Jeshion, R. (2013a). Expressivism and the Offensiveness of Slurs. *Philosophical Perspectives*, 27, 231-259.
- Jeshion, R. (2016). Slur Creation, Bigotry Formation: The Power of Expressivism. *Phenomenology and Mind*, 11, 130-139.
- Jeshion, R. (2013b). Slurs and Stereotypes. *Analytic Philosophy*, 54, 314-329.
- Jeshion, R. (2017). Loaded Words and Expressive Words: Assessing Two Semantic Frameworks for Slurs. *Croatian Journal of Philosophy*, 17, 111-130.
- Levinson, S. (2000). *Presumptive Meanings*. Cambridge, MA: MIT Press.
- Potts, C. (2007). The Expressive Dimension. *Theoretical Linguistics*, 33, 165-97.
- Popa-Wyatt M., & Wyatt J. (2017). Slurs, roles and power. *Philosophical Studies*, 175, 2879-2906.
- Richard, M. (2008). *When Truth Gives Out*. Oxford: Oxford University Press.
- Saka, P. (2007). *How to Think About Meaning*. Dordrecht: Springer.

- Tirrell, L. (1999). Derogatory Terms: Racism, Sexism and the Inferential Role Theory of Meaning. In (Eds.) K. Oliver, & C. Hendricks, *Language and Liberation: Feminism, Philosophy and Language*, Albany, NY: State University of New York Press.
- Vallée, R. (2014). Slurring and common knowledge of ordinary language. *Journal of Pragmatics*, 6, 78-90.
- Wattles, J. (2018). A Top Netflix Executive is out after using the N-Word, *CNN Money*, Cable News Network, 22 June 2018. Retrieved from: money.cnn.com/2018/06/22/news/companies/netflix-spokesperson-nword/index.html.
- Williamson, T. (2009). Reference, Inference, and the Semantics of Pejoratives. In J. Almog, & P. Leonardi (Eds.) *The Philosophy of David Kaplan* (pp. 137–158). Oxford: Oxford University Press.