Inferentialist semantics for lexicalized social meanings

Leopold Hess

Received: 31 August 2021 / Accepted: 14 July 2022 © The Author(s) 2022

Abstract

This paper offers a general model of the semantics of lexicalized social meanings, i.e. semiotic properties of certain expressions in a socio-political context. Examples include slurs, problematically charged expressions such as inner city, as well as terms such as mother, which also carry implicit ideological associations. Insofar as their linguistic properties are concerned, social meanings can be construed as context-structuring devices: without introducing specific at-issue contents, they evoke background assumptions which shape the context of conversation. An inferentialist model of discourse is developed to account for this effect, in which the discursive significance of an utterance is defined as the set of inferences it licenses relative to a discursive context. A discursive context is a set of propositions that can serve as auxiliary premises in material inferences, together with a salience ranking that makes some of these propositions more readily available and therefore more relevant to determining discursive significance. Social meanings are defined as functions on discursive contexts that modify the salience ranking, increasing the salience of certain assumptions and stereotypes. As a result, they impact the discursive significance of utterances indirectly and independently of at-issue contents. They are also largely independent of speaker intentions in virtue of the ideological nature of discursive contexts.

Keywords Social meaning · Inferentialism · Slurs · Dogwhistles · Discourse

1 Introduction

Philosophers of language have recently taken interest in social meaning, a concept whose origins can be traced to social philosophy and legal theory (cf. Kahan (1997), Haslanger (2014)). Social meanings in this sense are the conceptual schemas that
guide our interaction with the world and, especially, with each other. In other words, social meanings are semiotic properties that can attach to institutions, laws, practices, linguistic expressions etc. An example may be the social meaning of the institution of marriage (together with the laws that circumscribe it), which, as Ralph Wedgewood argues, “consists of the understandings and expectations regarding marriage that almost all members of society share [...] These assumptions seem to include the following: normally, marriage involves sexual intimacy [...] the couple’s cooperation in dealing with the domestic and economic necessities of life [...], and it is entered into with a mutual long-term commitment to sustaining the relationship.” This “meaning of marriage” can shape people’s understanding of romantic and familial relationships as well as practical choices in life (whether they are or want to be married or not).

My focus here is on social meanings insofar as they attach not to laws or practices, but to linguistic expressions (cf. Haslanger (2012), Stanley (2015), chap. 4) and do so in a relatively constant, conventionalized manner. This is what I will understand by lexicalized social meanings and the purpose of the present paper is to treat them as properly meanings, i.e. as a linguistic category. An example that has been widely debated in recent philosophical literature is the meaning of slurs: offensive epithets that refer to groups defined on the basis of some (perceived) demographic property, such as race, ethnicity, sexual orientation, religion etc., and convey both derogatory attitudes and harmful stereotypes concerning those groups and their standing in social hierarchies.

Slurs do not exhaust the category of lexicalized social meanings as I define it (and in fact my account will only partially concern them). It also includes other politically and morally charged vocabulary, which may be a vehicle of prejudicial or discriminatory contents without being overtly offensive, such as welfare, inner city, (illegal) alien etc. For instance, as Stanley (2015) argues, the use of inner city in American public discourse routinely conveys negative stereotypes about the predominantly Black inhabitants of downtown areas of US cities. In some sense—and it is my purpose here to elucidate this sense—any use of inner city carries the message that Blacks are lazy, prone to violence etc.  

Lexicalized social meanings may also be, or appear to be, more innocuous. Besides the institution of marriage, the term marriage itself also carries social meanings (as do the associated terms husband, wife, spouse, bride etc.). Other examples include words such as mother, woman, citizen, property. For example, Haslanger (2014) points out that mother carries the assumption that “one’s sex is relevant to one’s parental nurturing” (in this way it contrasts with ungendered parent, which is not to say that the latter term is not a vehicle of social meaning as well; see the discussion of parent vs guardian vs progenitor in Haslanger (2012), Chap. 14).

---

1 An arguably distinct but related concept of social meaning is employed in sociolinguistics (cf. Eckert, 2018; Hall-Lew et al., 2021) See Beltrama (2020) for a review of the recent surge of interest in social meaning in philosophical semantics and pragmatics.

2 In a column for New York Times (https://opinionator.blogs.nytimes.com/2012/05/24/marriage-meaning-and-equality, accessed on Aug 9th, 2021), cited in Stanley (2015).

3 While I will not mention any actual slurs in this paper, I cannot avoid citing some stereotypical statements and notions that may be deeply harmful. I assure the reader that none of them are hereby endorsed and apologize for any unintentional offense this may cause.

Springer
Social meanings, including lexicalized ones, have a variety of effects. They shape our practical reasoning and affective dispositions as we negotiate our way through the social world. But they also have strictly discursive effects: they shape our conversations and our understanding of what is said and debated, especially in public discourse. It is this latter kind of effect I will focus on here, and I will propose a semantic analysis of lexicalized social meanings based on a dynamic construal of Brandomian inferentialism (cf. (Brandom, 1994)). However, unlike inferentialists such as Brandom or Whiting (Whiting, 2008), I will not be concerned with the literal, explicit (at-issue) contents of expressions that carry social meanings, but only with these social meanings considered as a separate dimension. In Sect. 2, I expand the characterization of lexicalized social meanings and propose they should be treated as context-structuring devices. In Sect. 3, I clarify the relation between the semantics of social meaning I propose and theories of slurs, including a more familiar application of inferentialist semantics (going back to Michael Dummett’s famous analysis of Boche). Social meaning is only one aspect of the derogatory content of slurs, but it crucially distinguishes them from other kinds of pejoratives. This section concludes with a birds-eye view of the account to be developed. Section 4 introduces definitions of discursive context and discursive significance. In Sect. 5, I explicate the semantics of lexicalized social meanings in the resulting framework, as well as discuss the special case of “dogwhistles” and offer some remarks on the questions of the origin and evolution of social meanings (including reclamation of slurs). Section 6 expands on the more philosophical aspects of social meanings in discursive contexts, and especially their relation to ideology.

2 Social meanings as context-structuring devices

This section provides an initial characterization of what I dubbed lexicalized social meanings, based largely on the extensive discussion in Stanley (2015) as well as on observations commonly made in the literature concerning the derogatory meanings of slurs, which I extend (as does Stanley) to other kinds of social meaning-bearing expressions, such as the abovementioned welfare or mother. Focusing on slurs is useful because they tend to elicit strong and fairly stable intuitions regarding their semantic properties, and because they have been widely discussed in philosophy of language in the recent years. This is not to deny that there are important differences between slurs and expressions like welfare; I will discuss this in more detail in the next section, where I will propose that social meaning is only one aspect of the derogatory content of slurs.

The two basic properties of lexicalized social meanings are that they are not-at-issue contents and they are conveyed largely regardless of context and speaker intentions. (With respect to slurs the latter property is sometimes termed “derogatory autonomy”;

---

4 For a robust example, see Tali Mendelberg’s research on how racial stereotypes associated with terms like welfare or the poor in the US public discourse affect Americans’ political opinions.

5 The philosophical literature on slurs is already vast and rapidly growing; I will engage only with a small fragment of it here. For recent surveys, see Popa-Wyatt (2020); Hess (2021).

6 In the terminology introduced by Potts (2005). In the next section I will propose that they are more specifically implicit not-at-issue contents.
see Hom (2008), Anderson and Lepore (2013).) Let me briefly explicate both aspects and illustrate them with examples. Not-at-issue contents are independent from and supplemental with respect to the main (“at-issue”) content of an utterance; they are not part of what is asserted (if the utterance is an assertion), but are conveyed in addition to that. Examples include appositive and parenthetical clauses as well as expressive interjections or epithets, as seen in the following examples:

(1) That bastard Jake got promoted again.
(2) Jake got promoted again, I hear.
(3) Jake, the boss’s nephew, got promoted again.

In each case the at-issue content of the utterance is the proposition that Jake got promoted. “I hear”, “bastard”, “the boss’s nephew” convey supplemental material about the speaker’s attitude or epistemic state, or about Jake himself. This material is not asserted; it is not, in Stalnakerian terms, a proposal to update the common ground—which is not to say that it does not impact the common ground of a conversation. On the contrary, it may seem to be introduced directly without leaving room for negotiation or questioning—a hearer of (1) may challenge the assertion that Jake got promoted, but not the speaker’s negative attitude towards “that bastard”. Moreover, we should note that not-at-issue contents need not be unimportant asides or comments (as may be the case in (2), where “I hear” serves merely to hedge the speaker’s assertion), but can be an essential part of what is communicated. This seems to be the case in (1) and (3), where the speaker’s attitude towards Jake or the insinuation that his promotion is undeserved are very much the point of the utterance.

Importantly, unlike many other kinds of not-at-issue contents, social meanings may be carried by their associated expressions even if the speaker does not intend to convey them. The literature on slurs is rife with examples of such derogatory autonomy. Let S stand in for a slur for some group. Even if uttered sincerely, (4) does not fail to derogate that group:

(4) I have nothing against Ss. Some of my best friends are S.

Stanley (2015) discusses extensively how the problematic social meanings associated with the vocabulary concerning social benefits in the US, such as welfare, or terms such as inner city, make it difficult to engage in a reasonable debate about these topics without conveying prejudiced assumptions about race and poverty, even if one does not endorse any racist attitudes or practices. Similar observations are made by Elizabeth Anderson (2010, Chap. 3). Fraser and Gordon (1994) identify the same problem from a different perspective, foregrounding the sexist as well as racist dimensions of anti-welfare propaganda. The same can be said about Haslanger’s example of mother: anytime someone is described or referred to as a mother, certain assumptions about the relation between sex/gender and parental roles are implicitly, and often unwittingly, brought to bear on the conversation.

The combination of these two properties—not-at-issueness and independence of communicative intentions—results in the capacity of lexicalized social meanings to impact, in ways forceful or subtle, the shape of any conversation in which the expressions to which they attach are used. How can this impact be described? I propose, and

\[\text{See Murray (2014) for an account of not-at-issue content as directly updating the common ground.}\]
the rest of this paper will be devoted to substantiating this proposal, to construe it not as the introduction, negotiable or not, of any specific contents into the conversational context or common ground, but rather as structuring the context in certain ways.

With respect to slurs, a view like this is suggested by Davis and McCready (2020), who argue that the meaning of slurs, beyond expressing the attitudes or beliefs of a speaker, “serves to invoke and bring to contextual prominence a (possibly hidden) bundle of historical facts, prejudices, and social stereotypes about the group.” It is this invocation of structures of oppression that accounts for the derogatory force of slurs and distinguishes them from more mundane pejoratives. However, as forcefully argued by Stanley (2015), slurs should not be considered to be unique in this respect, as many words that are not overtly pejorative have a similar effect (I will qualify this claim and depart somewhat from his position in the next section). Stanley is also more explicit about the specifically discursive effects of social meanings, writing about the example of welfare:

When the news media connects images of urban Blacks repeatedly with mentions of the term “welfare”, the term “welfare” comes to have the not-at-issue content that Blacks are lazy. [...] This does not mean that someone hearing the term “welfare” automatically comes to believe that Blacks are lazy. It does mean that they may have to shift to different vocabulary, or consciously resist the effects of the association, in conversation or otherwise, to deter the propagandistic effect. (p. 138)

This makes democratic deliberation about the merits and problems with such programs more difficult; it requires first fighting about vocabulary. Raising doubts about such programs requires using the standard terminology for them. But the standard terminology affects the discussion by making salient these long-standing racial stereotypes. It therefore becomes difficult to criticize these programs without seeming to be a racist. (p. 161)

I cite these passages at length, because they make some points that are crucial for the account of social meanings I propose here (and they also hint at an explanation of how social meanings come to be, which is a question I will return to later). The effect of the not-at-issue content carried by words like welfare is not a matter of expressing beliefs by a speaker, or creating beliefs in the hearer, but of “affecting the discussion” by making salient problematic stereotypes. As a result, the discourse is shaped in a certain way regardless of the interlocutors’ attitudes and intentions. Even if they do not agree with them, racist stereotypes are now a salient part of the conversational background.

Stanley considers several mechanisms that could account for the discursive effects of social meanings, all of which focus on how they can erode empathy for targeted groups and undermine the reasonableness of political debate. This is not my focus, so I will not discuss his proposals in any detail. Instead, I will offer a more general model of context-structuring devices, which can also accommodate the social meanings of words such as mother or parent. The latter may have effects both more subtle and more complex than shaping how we empathize with certain groups of people, but they
also structure the way we think and talk about topics such as parenting, gender roles etc.\(^8\) (And similarly for other examples such as refugee, property etc.)

To summarize this section, I consider lexicalized social meanings, conveyed in addition to whatever literal or at-issue contents their vehicle expressions have, as context-structuring devices, by which I understand that the effect they have on a conversation is to shape its context in a way that brings certain stereotypical assumptions to salience and thereby to impact the meaning of utterances, whether or not this is intended by interlocutors. To account for this effect, I develop a discourse model incorporating crucial insights of Brandomian inferentialism (although without committing to inferentialism as a theory of literal meanings). This will be developed in detail in Sects. 4 and 5, but first a closer look at slurs in their role as social meaning vehicles is required.

### 3 Slurs, pejoratives, and social meanings

The purpose of this section is to position my proposed account with respect to existing theories of slurs. I take slurs to be an important example of expressions carrying social meanings, but (unlike Stanley) I do not assume that there is no difference between slurs and words such as welfare. In Sect. 3.1 I propose to understand slurs as having both a social meaning and an expressive meaning. The latter kind of meaning is shared by slurs with other pejoratives; the former meaning is the only one I am concerned with. This will also lead me, in 3.2, to distinguish my approach from a more familiar discussion of slurs and inferentialist semantics. Having done all this necessary ground-clearing work, in Sect. 3.3 I outline the account of social meanings that I will develop in the rest of this paper.

#### 3.1 Are slurs special?

Stanley (2015) argues that slurs are not ‘special’, that is, not only are they not unique in carrying discriminatory social meanings—the kind of meaning conveyed by, say, an anti-Black racist slur is not importantly different from the implicitly racist meaning of welfare. This claim has been criticized among others by Cepollaro and Torrengo (2018), who point out salient differences between slurs and words like welfare (these differences should be even more salient if we consider social meaning-carrying expressions such as mother). In response, Stanley (2018) clarified his position by allowing that there are differences in degree between these categories of expressions. I will offer a different response which preserves the intuition that slurs are special in some way while supporting an account of their social meanings as continuous with non-slurring vocabulary.

Cepollaro and Torrengo point out two important differences between slurs and words like welfare. First, it is nearly impossible to use a slur without conveying its

---

\(^8\) See the discussion in the already cited essays by Haslanger (2012; 2014). Note that Haslanger’s concern is with social structures more broadly, while I only focus on linguistic effects of social meanings.
derogatory meaning, but the social meanings of words like welfare are not conveyed in all contexts. They give the following example:

(5) Eisenhower took his turn with the interstate highway system and the creation of the department of health, education and welfare.

An utterance of (5) does not seem to invoke any racist stereotypes. Secondly, the racist contents associated with welfare and such are arguably implicit, while slurs are explicitly derogatory.

To these two differences I would add a third one: slurs convey strongly emotional meanings. There is a certain expressive load carried by any use of a racist slur that does not accompany even the most brazenly propagandistic uses of welfare or inner city—even if it is not intended by a speaker, it can certainly be felt by the hearers, and especially targets. (See Rappaport (2020) on what he calls the “toxicity” of slurs.)

Stanley (2018) responds to this criticism by granting that there are differences between slurs and non-slur expressions, but that they are only differences in degree. He points to examples of reclaimed slurs that offer counter-examples to Cepollaro and Torreno’s claims that slurs (unlike welfare etc.) always convey derogatory meanings. The value of these counterexamples, however, depends on one’s theory of reclamation. Stanley and his critics differ strongly on this point (Cepollaro and Torreno treat reclaimed slurs as cases of polysemy or non-literal use). I will not attempt to adjudicate this dispute. There’s a more basic problem with Stanley’s response. While it is not implausible that the difference between slurs and other lexicalized social meanings is one of degree, rather than kind, the question remains: degree of what? Simply positing such a difference does not yet provide an explanation.

Taking a different tack, I propose to distinguish two aspects of the derogatory, non-truth-conditional/not-at-issue content of slurs: an expressive meaning and a social meaning. I interpret Davis and McCready, as cited in the previous section, to suggest something like that (although without making an explicit distinction): the use of a slur expresses a speaker’s attitude and “brings to contextual prominence a […] bundle of historical facts, prejudices and social stereotypes”.

There are independent reasons to make such a distinction. Many authors (e.g. Potts (2007), Whiting (2008), Jeshion (2013), McCready (2010)) treat slurs as having a two-dimensional semantics, with an at-issue, truth-conditional component that is identical for the slur and its neutral counterpart (i.e. an innocent term of reference for the group that the slur targets) and an attitudinal or expressive component. The latter component is a meaning of the same kind as that of pejoratives such as bastard or jerk: one that

---

9 One could quibble with that. In this particular instance, “welfare” appears as part of the name of a government agency, so it’s not clear if it is actually used rather than merely mentioned. And if it is, it’s not unlikely that it would result in bringing to salience some stereotypical assumptions concerning welfare and its recipients, depending on how the conversation would continue. Nevertheless, it can be granted that in most cases a slur would convey derogation regardless of such considerations of conversational context.

10 Stanley acknowledges this explicitly and points to forthcoming work (a book co-authored with David Beaver) in which such an explanation is to be provided. This work has not been published yet at the time of my writing, and in any case it is likely to entail an account of social meanings different from the one I propose here.
is expressive of a negative attitude (which can be described as contempt, hatred etc. or left unspecified).\textsuperscript{11}

There are, however, obvious differences between, say, a racist slur, and even a quite intense pejorative like \textit{motherfucker}. Slurs effect a distinct kind of offense, not only to those who they target but also to bystanders. Their use is usually considered to be not just impolite (as with swearwords) but morally reproachable and politically problematic.\textsuperscript{12}

Some authors, e.g. Bach (2018) and Jeshion (2013), explicitly admit that their theories are not meant to account for the difference between slurs and other pejoratives (others seem happy to ignore this issue). Jeshion observes that whether or not certain terms are considered slurs depends on whether or not the derogation of their targets is considered appropriate or permitted. Bach suggests that his semantics can be supplemented by a pragmatic account along the lines of Nunberg (2018) or Bolinger (2017), who focus on the social practices and institutions in which uses of slurs are embedded. In this spirit, I will posit that the difference between slurs and other pejoratives is their \textit{social meaning}.

On the view I am proposing, there are accordingly three distinguishable components to the meaning of a slur:

1. an at-issue meaning which secures reference;
2. an attitudinal meaning which effects pejoration;
3. a social meaning which has the discursive effects described in Sect. 2.

In this paper I have nothing more to say about components 1 and 2 and the account of social meaning I develop is intended to be compatible with a broad range of theories. In particular, component 1 can be modelled truth-conditionally (as most authors do) or in inferentialist terms (cf. Whiting (2008) and see the next subsection). Component 2 can be conveyed as an implicature or a presupposition and it can be expressive or descriptive (or perhaps it can be described in some yet other way, but these are the main options in the literature). In any case, I will assume that it does not distinguish slurs from at least some other pejorative terms that do not have the same social significance. There may be differences in degree, of course, but not in kind. What does distinguish them is component 3, the social meaning—which in turn is of the same kind as the social meanings of \textit{welfare, mother} etc.

In other words, on the view I am proposing, slurs belong to the intersection of two lexical categories: pejoratives and lexicalized social meanings. They are “special” in this way: the pejorative meaning accounts for the differences mentioned earlier—slurs are explicitly offensive, cannot be used in a neutral way and carry an emotional load in the same way that strong insults such as \textit{motherfucker} do. But at the same

\textsuperscript{11} Note that such assimilation does not require a specifically \textit{expressivist} account. Bach (2018) proposes descriptive, evaluative meanings for slurs (predicating contemptibility rather than expressing contempt) but also suggests that they are similar to the meanings of pejoratives like \textit{asshole}. Cepollaro and Stojanovic (2016) go in a somewhat different direction, offering a unified two-dimensional account of slurs and thick terms such as \textit{generous}. Even if some of these terms, like \textit{lewd}, may be considered objectionable by many speakers, they are clearly not as offensive as slurs.

\textsuperscript{12} Nunberg (2018) discusses this point extensively and based on it argues against semantic approaches to slurs, instead treating them as having only a neutral at-issue content plus a social meaning. But that means that the problems pointed out by Cepollaro and Torrengo apply to Nunberg as much as to Stanley.
time in one aspect they are continuous with the kind of vocabulary that Stanley is most interested in: unlike *motherfucker* and like *welfare*, they invoke a background of bigoted stereotypes and insinuations. It is only this aspect that is relevant to my present purposes.

A terminological note is in order. Both component 2 and 3 of the meaning of slurs can be considered to be not-at-issue. What distinguishes them is that the former is explicit while the latter implicit. In what follows I will no longer use the labels at-issue vs not-at-issue, but distinguish social meanings as implicit in opposition to explicit literal meanings.  

### 3.2 Pejoratives and inferentialism

As I have hopefully made clear in the previous subsection, my semantics for social meanings applies to slurs but is not intended to provide a comprehensive account of their meaning. Accordingly, even though I develop an inferentialist account, it should not be considered to be part of a long-standing debate about the applicability of inferentialist semantics to slurs. This debate, as I will now briefly explain, concerns the *explicit* meanings attributed to slurs, not their implicit social meanings.

An inferentialist construal of the meaning of derogatory epithets was first proposed by Dummett (1973), and has been endorsed at least to some extent by i.a. Brandom (2000) and Tirrell (1999, 2017). Dummett’s example was *Boche*, by now an outdated pejorative term for Germans, the associations of which imply that members of this nation are prone to cruelty and barbarism. Within a semantic theory that defines meanings of expressions through rules governing their use in inferences, the connection between descriptive and pejorative aspects of *Boche* can be captured through rules of something like the following shape:  

\[
\text{Boche-introduction:} \quad \frac{x \text{ is German}}{x \text{ is Boche}}
\]

\[
\text{Boche-elimination:} \quad \frac{x \text{ is Boche}}{x \text{ is cruel}}
\]

However, these rules are importantly flawed from the perspective of the inferentialist theory. There is a mismatch, a “disharmony”, between the introduction and the elimination rule. As a result, they allow for inferences from “x is German” to “x is cruel”. For inferentialists such as Dummett and Brandom this is an important part of the explanation of what is *wrong* with slurs: their use licenses inferences that no reasonable speaker should endorse.

This account of the meaning of slurs has met with significant criticism from Hornsby (2001) and especially Williamson (2009). Williamson argues that the consequences of application of a slur cannot be explicitly spelled out as they would need to be to formulate a determinate elimination rule (because it’s difficult to say what exactly a

---

13 This classification assumes that there are no implicit at-issue meanings. I take no stance on this, but if the reader finds a reason to posit such meanings, my uses of “implicit” can be understood as short for “implicit not-at-issue”.

14 This is Williamson’s (2009) interpretation of Dummett’s informal remarks.

15 Note that Brandom, unlike Dummett, does not endorse the formal constraint of harmony, but argues that inferences licensed by *Boche* and the like are materially unacceptable.
slur conveys about its targets in descriptive terms); that it is impossible to determine the reference of a slur, given the disharmony between introduction and elimination rules (does Boche refer to Germans or to people who are cruel? There should be a definite answer at least for speakers who endorse the use of this word); and finally, that it is hard to say what understanding the slur consists in for speakers who do not endorse the flawed inference schema.

Others, e.g. Whiting (2008, 2013), have defended an inferentialist approach to slurs and the debate is ongoing. For its most recent developments, see Valtonen (2019) and Diaz-Leon (2020), who propose different modifications of the specification of inferential rules for slurs, to avoid criticisms like those mounted by Williamson.

I do not discuss this in any more detail as my purpose here is merely to note that, from the perspective outlined in the previous subsection, this debate focuses on the applicability of inferentialism to the explicit meaning of slurs (which Whiting construes in a two-dimensional way as a combination of at-issue content modeled in an inferentialist fashion and a non-cognitive attitudinal content; earlier authors did not make such a distinction). In this context, social meaning is not specifically distinguished. On the contrary, my focus is on implicit social meanings. On my view, “Hans is a Boche” may be understood to assert (at-issue) that Hans is German and (explicitly not-at-issue) express contempt for him or predicate contemptibility of him, or Germans in general—and also to (implicitly not-at-issue) invoke negative stereotypes about Germans, such as that they are especially cruel etc. It is this last aspect that I focus on.

### 3.3 Overview of the semantics for social meanings

So far I have clarified what I understand by “social meaning” and how it relates to slurs and some familiar theories of their meaning. In the rest of this paper I will develop my positive account. For the reader’s convenience, I first outline its main points.

A discursive significance of an utterance can be defined as the set of inferences one can draw from the content of that utterance. This is relative to what auxiliary premises are available in a given context. Availability of premises, and therefore discursive significance, is impacted by salience. Social meanings are semantic devices that shape discursive contexts by making the stereotypical assumptions associated with them salient as auxiliary premises. Thus, the significance of an utterance containing some lexicalized social meaning comes to include those stereotypical assumptions and what can be inferred from them.

A special kind of lexicalized social meanings are so-called dogwhistles, i.e. expressions that are intended to carry additional implicit meanings to only one segment of the audience. On my account this is naturally modeled in terms of an utterance having different discursive significance for different segments of the audience in light of what auxiliary premises are available to be made salient for them.

---

16 Thus, an inferentialist account of the explicit meaning of slurs, if it is tenable, could be combined with an inferentialist account of the social meaning, but this would still be two distinct models of two different meaning components.
My account is not intended to explain the origin of social meanings, but following Stanley I accept a rough picture on which they arise due to a stable association between the uses of a certain expression and the presence of certain stereotypes. This effect may be stronger in the case of slurs in virtue of their attitudinal content. On the other hand, in the case of reclaimed slurs, a change to social meaning may bring about a change to expressive meaning.

These points are developed in Sects. 4 and 5. In Sect. 6, I conclude with some broader reflections on meaning and ideology.

4 Discursive significance

I will now introduce the notions of discursive significance and discursive context (loosely based on Brandom (1994)); in the next section I will provide an analysis of lexicalized social meanings as updates to discursive contexts.

The discursive significance $ of an utterance (for simplicity’s sake I focus only on assertions, but the context-structuring effects of social meanings described later can occur with other speech acts as well) is the set of inferences it licenses in a given discursive context. The discursive context $\Delta$ can be defined in a first approximation (soon to be revised) as a set of propositions that can be used as auxiliary premises for inferences. For now we can assume that these propositions are common ground (i.e. mutually accepted) between all interlocutors, although this requirement will be later relaxed. Schematically:

$$
\text{Discursive significance: } S(U)^\Delta = p: U \rightsquigarrow_\Delta p,
$$

where $U$ is an utterance, $p$ a proposition variable, and $\rightsquigarrow$ is a material (defeasible, nonmonotonic) inference relation.\(^{17}\)

The best way to explicate these notions is through examples:

(6) It’s going to rain soon.

(7) Benjamin Franklin was born in 1706.

If Ali says (6) to Brianna, and it is part of the discursive context (it’s common ground between them) that Brianna is going out soon, the discursive significance of this utterance may be that Brianna should take an umbrella. That is, the proposition that Brianna should take an umbrella can be inferred from the premise that it is going to rain soon combined with the premise that Brianna is going out. In the absence of the latter, (6) would not have this discursive significance. Moreover, the inference is nonmonotonic (or defeasible): if it is also part of the discursive context that Brianna does not mind the rain, or that she is taking a taxi, the proposition that she should take an umbrella will not be part of the inferences licensed by “It’s going to rain soon”.

The discursive significance of (7) may be that the first Postmaster General of the United States was born in 1706, if the discursive context contains the auxiliary premise that Benjamin Franklin was the first Postmaster General. It may also be that the inventor of the bifocals was born in 1706, if the discursive context contains the proposition

\(^{17}\) I follow Brandom’s understanding of material inferences, but nothing in my argument regarding social meanings hinges on it.
that Franklin was the inventor of the bifocals. It may of course contain both these propositions. The discursive significance of (7) may even be that that the author of Waverley was born in 1706, if the context provides the auxiliary premise that Benjamin Franklin was the author of Waverley. There is no reason to assume that the discursive context contains only true propositions—this is a crucial point for the analysis of social meanings which are often associated with false assumptions and stereotypes.

Discursive significance in my understanding is an intentionally broad category describing the total conversational impact of an utterance, and thereby covering many things that would traditionally be regimented as either entailments (as in (7)) or conversational implicatures (as in (6)) and possibly other categories like informative presuppositions.\textsuperscript{18} It could be objected that it is too broad to be of analytical value, but I should emphasize that it is not intended as an explanatory concept as such—it is the more nuanced aspects of discursive significance discussed below that will be important for analyzing lexicalized social meaning. For this reason I also do not use the term “inferential significance” introduced for what I take to be the same concept by Drobňák (2020) and Kaluziński (2020), who employ it to distinguish and analyze sentence meaning and utterance meaning. This is not relevant to my purposes here, but a reader who wishes to know how discursive/inferential significance can be spelled out in more detail is referred to their work.

What is relevant to my purposes is that to account for variations in discursive significance we should treat the discursive context not simply as a set of propositions akin to a traditional concept of common ground, but as a set that is ranked for salience (in Sects. 5.1 and 6 I further revise this understanding of what discursive contexts are). The importance of salience ranking can be illustrated with the following example

(8) Tom is 15 years old.

If Charlize says (8) to Deepak, its discursive significance may be that Tom is old, if the discursive context contains the additional premise that Tom is a cat. It will not have this significance, quite the contrary, if the context contains the premise that Tom is a (human) person. Schematically again:

\[
\begin{align*}
S(U)^{\Delta} \ni p, & \text{ if } \Delta \text{ contains } q = \text{Tom is a cat} \\
S(U)^{\Delta} \ni \neg p, & \text{ if } \Delta \text{ contains } r = \text{Tom is a human}
\end{align*}
\]

But it may very well happen that there is more than one Tom mutually known to Charlize and Deepak, one of them perhaps being Charlize’s cat, and another Deepak’s brother. In this case, \(\Delta\) contains both \(q\) and \(r\). However, the discursive significance of (8) should not be taken to contain both the propositions that Tom is old and that Tom is young. Even though these concern different Toms, so there is no contradiction, pragmatics of a conversation dictate that only one should be the relevant uptake of Charlize’s utterance. For this reason, the propositions in the discursive context should be ranked for salience and only ones that are salient enough will be available as

\textsuperscript{18} Note that discursive significance concerns what can be inferred by the hearer. This contrasts with concepts concerning the speaker’s commitments, such as conventional implicature. As pointed out by an anonymous reviewer, this is another way in which my approach differs from more standard theories of slurs (cf. McCready (2010); Whiting (2013)).
auxiliary premises that support an utterance’s full significance. The context should then be represented as a structure:

$$\Delta = \{D, \rho\}$$

where \(D\) is a set of propositions and \(\rho\) is a salience ranking over \(D\).

The salience of propositions in \(D\) may be affected by previous utterances, in ways both explicit and implicit. Charlize may have said “My cat’s name is Tom”, thereby directly increasing the salience of Tom the cat and propositions concerning him over any other Toms known to her and Deepak. Or she could say “I’m worried about Tom’s health. He’s 15 already.”, which would indirectly raise the salience of Tom the cat, as 15 is usually not old enough for humans to be worried about their health. It could also be the case that all other Toms mutually known to Charlize and Deepak are adults, and so the discursive context contains the propositions that Tom, Deepak’s brother, is 38; Tom the bartender is 27; etc. None of these could serve as auxiliary premises to be combined with the proposition that Tom is 15, on pain of contradiction, and so the utterance of (8) would automatically make Tom the cat the only possible object of reference, and the proposition that Tom is a cat the only salient auxiliary premise.

Consider another example:

(9) Viktor is very short.

Let’s say, for simplicity’s sake, that an adult man is very short if he is shorter than 165 cm, unless he is a basketball player, in which case he is very short (for a basketball player) if shorter than 190 cm. Presumably, these two propositions—that men in general count as very short if shorter than 165 cm and that male basketball players count as very short if shorter than 190 cm—are part of a cultural background and can be available as auxiliary premises in a discursive context in any conversation between people sharing this background. But in a particular conversation, one of them may be more salient than the other (even if we’re speaking about the same Viktor). If the conversation concerns basketball, the discursive significance of (8) will only be that Viktor is shorter than 190 cm. Otherwise, it may be that he is even shorter than that, under 165 cm.

The purpose of these examples was not to analyze in detail the various ways in which salience of potential auxiliary premises in a discursive context can be manipulated (there is no need for it here), but only to emphasize that a salience ranking is important to characterizing the discursive significance, i.e. the total conversational impact, of an utterance relative to a context. In the next section I return to lexicalized social meanings and propose to understand their context-structuring function as operating not on what propositions are included in the set \(D\) of the discursive context (the counterpart of common ground) but on their salience ranking (\(\rho\)).

### 5 Social meanings and discursive significance

On the account I propose, the discursive effect of lexicalized social meanings is to structure the context of a conversation through modifying the salience ranking (\(\rho\)) of propositions that can serve as auxiliary premises (\(D\)) to constitute the discursive signif-
icance of utterances employing expressions to which those social meanings attach. The modified ranking awards greater salience to stereotypical assumptions associated with the given social meaning. This is a way of substantiating the notion presented in Sect. 2 that the effect of lexicalized social meanings is to evoke background assumptions that influence the interpretation of utterances.

Thus, the social meaning of an expression E can be defined as an update function M on discursive contexts:

\[ \{D, \rho\} \oplus M = \{D, \rho'\} \]

where \( \rho' \) updates \( \rho \) by increasing the salience of stereotypes associated with E.

What assumptions exactly become salient in this way is a matter of the broader social context: the practices, institutions and ideologies that shape a society’s understanding of the relevant categories such as race, gender, romantic relationships, citizenship, property etc. etc.

This relation between lexicalized social meanings and the socio-political context can be construed on the model of Hom’s (2008, 2012) combinatorial externalism, originally developed as a theory of slurs. According to Hom, the semantic value of a slur “is a complex, normative property of the form: ought to be subject to such-and-such discriminatory practices for having such-and-such stereotypical properties all because of belonging to such-and-such group.” In the case of e.g. Boche,\(^{19}\) this could be something like ‘ought to be treated as an enemy and denied basic personal rights for being exceptionally cruel and barbaric because of belonging to the German nation’. Combinatorial externalism posits that such semantic values are derived from and supported by institutions and practices of prejudice and discrimination.

The crucial difference between Hom’s account and mine is that Hom treats the social meanings attached to slurs as part of their literal, at-issue content. This is implausible and gives rise to wrong predictions concerning the projective behavior of slurs, as pointed out i.a. by Sennet and Copp (2015) and Cepollaro and Thommen (2019). My account avoids these problems by taking on board the externalist claim, but construing social meanings not as literal contents but as context-structuring devices that raise the salience of background assumptions concerning the slur’s targets. The complex meaning of Boche proposed above is evoked by the slur, but it is not part of the literal, explicit content of “Hans is a Boche”. As I explained in Sect. 3.1 I consider the explicit content to be some combination of a descriptive proposition that Hans is German and an attitudinal meaning with a negative valence towards Germans, but I intend my account of social meaning to be compatible with many different accounts of the full meaning of slurs. The only aspect I consider here are the social meanings that attach to expressions in addition to their literal contents.

Social meanings thus construed are not unique to slurs and so Hom’s combinatorial externalism can be applied to a broader range of expressions, but only if this distinction between explicit and implicit contents is made. It would be even less plausible to construe racist stereotypes or assumptions about gender roles as part of the literal

\(^{19}\) I use Boche as the only example of a slur here, as it is seems to be archaic enough to be void of any offensiveness, but it should be understood to stand in for slurs and harmful stereotypes targeting groups actually facing bigotry and oppression in present-day society.
content of welfare or mother. Instead, uses of these expressions can be interpreted like Boche was in the preceding paragraphs. Thus, an utterance of “We need to cut welfare” does not explicitly say that Blacks are lazy or anything similar, but it raises the salience of such assumptions in the discourse context. And so the discursive significance of such an utterance may be (may contain the proposition) that the reason why welfare programs should be scaled back is because they undeservedly benefit Black people. Moreover, if Stanley and other critics of American political discourse are right, an utterance of “Welfare programs should be expanded” may, because of the social meaning attached to welfare, have a similar effect on the discursive context: raising the salience of anti-Black stereotypes as auxiliary premises in the discursive context. And so the significance of this utterance may be that welfare should be expanded because Black people cannot take care of themselves. In either case, assertions ostensibly concerning government spending acquire malicious overtones, distorting any conversation about these topics regardless of the interlocutor’s own intentions and beliefs (unless the problem is explicitly addressed and racist assumptions are challenged).  

Similarly for mother and the like: any reference to someone as a mother, in virtue of the social meaning attached to this term, will increase the salience of assumptions concerning the relevance of one’s sex to one’s parental nurturing in the discourse. Thus, “Elfi is a mother” may have the significance of implying that Elfi should behave in a way according with her role as a female parent—without asserting any such proposition.

To summarize, I have outlined a dynamic, discourse-centered model of the semantics of lexicalized social meanings in which they are represented as functions on discursive contexts. By raising the salience of certain background assumptions that are candidate auxiliary premises, social meanings indirectly contribute to the determination of the discursive significance of utterances in which they are used. The most important aspect to emphasize is that social meanings are not at-issue or explicit contents and they remain in no direct relationship to at-issue contents of the expressions to which they attach. It is the broader socio-political context and, most importantly, the history of use that makes certain words carry certain social meanings. There are, of course, indirect relations between at-issue contents and social meanings—it is not an accident that a pejorative used to refer to, say, Jewish people carries social meanings related to anti-Semitism (I come back to this point in Sect. 5.2). But expressions that are equivalent in their literal, at-issue content may be associated with different social meanings. Downtown area does not seem to evoke the same racist stereotypes as inner city; the significance of parent (e.g. on a form) is different than that of father or mother. Social meanings and discursive contexts are interfaces between language and the extra-linguistic, social context.

This concludes the main part of my argument. In the next subsection I discuss how this general framework can be adapted to an interesting special case, and in the following one I offer a very rough sketch of the origin and evolution of social meanings.

---

20 See Saul (2018) on the possibility of undermining the effects of implicit racist messaging, somewhat underplayed by Stanley.
5.1 Dogwhistles

So far I have assumed that a discursive context is common to all interlocutors in a given conversation, so that the set $D$ can be identified with a set of common ground propositions, and the ranking function on them is also the same for everyone. However, this assumption can be relaxed and we can treat discursive significance as relative to individual interlocutors. This follows naturally from understanding discursive significance as determined by the availability of auxiliary premises. The latter may differ for interlocutors: even in one and the same conversation in which e.g. (6) is uttered (“Benjamin Franklin was born in 1706”), it may mean that the inventor of the bifocals was born in 1706 to one person and that the author of the *Waverley* was born in 1706 to another, depending on what additional premises they accept concerning Franklin.

This relativization of discursive significance is not an essential part of the general theory I propose here, but it makes it possible to account straightforwardly for a special case of lexicalized social meanings: so-called dogwhistles. In common parlance, dogwhistles are expressions that may appear “ordinary” and innocuous to the general audience, while communicating additional (often ideologically charged) meanings to those in a position to recognize them, in particular to those who share the speaker’s politics. Here is one oft-discussed example, from George W. Bush’s 2003 State of the Union Speech:

(10) Yet there’s power, wonder-working power, in the goodness and idealism and faith of the American people.

For people who don’t know better, “wonder-working power” could seem to be nothing more than a rhetorical flourish, as lofty and as empty as most political speech. In fact, however, it is a reference to a hymn popular among fundamentalist Christians, whose favor Bush intended to curry, without at the same time making his allegiance to their far-right politics openly recognizable. For this subset of the audience, “wonder-working power” connotes Christian faith, which makes salient an interpretation of Bush’s speech that goes beyond typical political boilerplate. In the model I present here, this interpretation can be reconstructed by assuming that for the fundamentalist Christians in the audience the discursive context in which (10) was interpreted contained a background premise such as “A good society is based on devotion to Christ”. The social meaning carried by “wonder-working power” modified this context by greatly increasing the salience of that premise. As a result, the discursive significance of Bush’s utterance included the proposition that the goodness and idealism and faith of the American people are inspired and sustained by their devotion to Christ. (These are of course simplifications, but they illustrate the basic mechanism.) At the same time, Bush could plausibly deny that this is what he was saying—not only because he did not directly assert anything of the sort, but also because for the general audi-

---

21 Cited in Saul (2018). Saul herself has a broader conception of dogwhistles that covers many, but not all, of what I defined as lexicalized social meanings—in her terminology, the standard understanding corresponds to the subcategory of *intentional overt dogwhistles*. There is also some debate concerning the nature of dogwhistles’ meaning (cf. Henderson and McCready (2018)). There may be reason to prefer a different definition and explanation of dogwhistles, but my aim here is merely to show that the general framework of inferential semantics for social meanings can be flexibly adapted to special cases.
ence, lacking the relevant background beliefs that could act as auxiliary premises, the utterance just did not have this significance.

Speaking more generally, a dogwhistle can be understood as a vehicle of lexicalized social meaning in the same way as any of the other examples discussed here: the social meaning structures discursive contexts in which it is evoked by raising the salience of certain associated background premises. What makes dogwhistles special is just the fact that their significance is not generally recognizable by everyone, because not everyone has the beliefs or awareness that would make the relevant premises available at all. While for the “insider” audience the social meaning will effect an important change to the discursive context, imbuing a given utterance with special significance, for the general audience it will act as an identity function, making no change to the context and leaving them to interpret the significance of the utterance based solely on its explicit content.

5.2 Origin and evolution of social meanings

What I propose in this paper is a semantic model of lexicalized social meaning, which, being a strictly synchronic theory, is not intended to explain where social meanings come from or how they develop. Nevertheless, it will be useful to sketch what such an explanation could look like, in the broadest strokes, if only to show that the kind of meaning I take social meanings to be are not too mysterious.

Here again I follow Stanley (2015, chapter 4) who suggests that a word like *welfare* comes to have the social meaning it has through regular co-occurrence of its use and the communication or insinuation of racist propositions. That is, through some period of time the word *welfare* was constantly used in public discourse in utterances that were either explicitly racist or communicated something racist implicitly through other already established means. As a result, a stable association has formed between this word and racist contents, so that the public is now primed to have those contents in mind whenever the word *welfare* is used, even if neither the speaker nor the hearers are ready to endorse them. These contents color the meaning of any utterance mentioning welfare, because, in my terminology, they co-determine the discursive significance of such utterances by serving as salient auxiliary premises.

This simple associationist picture glosses over many details which I will not attempt to fill in (and they are probably best left to social scientists and psychologists, rather than armchair philosophers), but even in its rough form it is useful for several purposes. First, it serves to explain why social meanings should attach to specific linguistic forms, i.e. why words that are more or less equivalent in explicit content do not share the same social meanings (e.g. *downtown* vs *inner city*). Such differences can be assumed to track histories of usage—if *inner city* is a word that has regularly been used in utterances with bigoted contents, and *downtown* is not, such contents will be associated with the former, but not the latter.

Second, with this picture we can easily understand why the social meanings of slurs should be more difficult to neutralize, displace or detach than in the case of words like *welfare*. For slurs are not only vehicles of social meanings, but also pejoratives. They
are typically used to directly insult their targets.\textsuperscript{22} And this can make the association between a word and its social meaning that much stronger, because someone who is ready to use insulting language about some group of people seems more likely to also make some bigoted assumptions about that group of people.

Third, this picture allows for the possibility of active contestation of social meanings and even reclamation of slurs, while showing that it may be very difficult. Stable associations that developed over long histories of use can be changed or decoupled, but that requires a lot of effort. An observation frequently made in the literature on slurs is that even with manifestly best intentions, individual attempts at using a slur in a way that is not derogatory may easily misfire. Anderson (2018) and Hess (2020) argue that successful reclamation presupposes the existence of social practices that attach a non-derogatory meaning to a slur. If robust and successful enough, such practices may even serve to change or neutralize the attitudinal meaning of a slur, if the association with non-bigoted social meanings becomes strong enough (consider the case of queer in “Queer Studies” etc.). Similarly, for welfare to lose its association with racist stereotypes, it is not enough for individual speakers to try and use the word without any such assumptions—a new form of discourse about welfare would have to develop.

How to fill in this associationist picture with details, and whether it is in the end the correct one are open questions, but for the present purposes it should suffice to show that the theory of social meanings I propose is consistent with some fairly commonsensical ideas about public discourse.

\section*{6 Social meanings and ideology}

I have suggested that the discursive context, or more specifically the set of propositions $D$ that is a part of it, should be understood along the lines of a standard notion of common ground. This is not entirely adequate for the purpose of analysing the context-structuring effects of social meanings. The use of expressions such as welfare or mother on my analysis increases the salience of certain propositions in the discursive context, but it is not plausible to assume that the relevant propositions are always present, even as background assumptions of initially low salience, in the common ground of any conversation. For the common ground contains propositions that are mutually believed or at least accepted in some weaker sense by the interlocutors, while lexicalized social meanings operate on propositions that need not be accepted by everyone. Not everyone accepts, even in a weakest sense, the stereotypes concerning ethnic groups and gender roles that were invoked in my discussion of examples in Sect. 5.

The set $D$ should therefore be considered to be broader than the set of common ground propositions. It also includes stereotypes, ideas and assumptions that are current in society, even if endorsed only by some members of it, as long as they are widely recognized as current, at least implicitly. Anderson (2010, p. 53-54) describes certain

\textsuperscript{22} See Bolinger (2017) on the centrality of insulting uses to the pragmatics of slurs.
racial stereotypes as having a *public standing*. A proposition has public standing\textsuperscript{23} between interlocutors if each of them entertains this proposition and it is a matter of mutual knowledge that each of them entertains this proposition. What is crucial is that entertaining, Anderson argues, does not require endorsing—one can simply have a certain proposition in mind. An example that she discusses is the stereotype of Black men as being especially prone to be violent criminals. As she observes, “[t]hat this is a prevailing stereotype is a matter of common knowledge among Americans.” It seems likely that such a stereotype is part of the social meaning of *inner city* (or in any case let us assume that for the sake of argument), that is, it is brought to contextual salience whenever this term is used. If so, then on the model I propose, this stereotype, and various inferences that can be based on it, will be part of the discursive significance of any utterance containing the term *inner city*, even if none of the interlocutors would be ready to endorse it. Public standing of social meanings is sufficient for their operation. (For terms such as *mother* or *woman* the effect may be less transparent, but is more pervasive.)

In other words, it need not be common ground that “Blacks are lazy” or “Sex is relevant to one’s parental nurturing”, but it is common ground for all (or most) members of a society that people may think so—and therefore such assumptions may be available as auxiliary premises for determining the discursive significance of utterances. Their default salience may be very low—especially in the case of the most bigoted and harmful stereotypes—but can be raised with the use of the kind of vocabulary that was the subject of my discussion here. Moreover, as the analysis of dogwhistles in the previous section shows, not even that much needs to be true. For some social meanings it is enough that they have public standing among a certain part of society, while the rest may be completely unaware of them.

Another way to put this is that discursive contexts are *ideological* structures. Ideology is notoriously defined in many, sometimes incompatible, ways, but here I will follow Swanson’s (ms) conception of ideology as a cluster “of mutually supporting beliefs, interests, norms, practices, values, affective dispositions, and ways of interpreting and interacting with the world.” Social meanings, as I defined them in Sect. 1, are elements of ideology thus understood. More specifically, with respect to the semantic model I propose, the role of ideology is in shaping discursive contexts (which in turn determine the discursive significance of utterances): it provides the contents of the set $D$, as well as the ranking $\rho$. The proposition that one’s sex is relevant to one’s parental nurturing is available as an auxiliary premise, the salience of which can be raised with the use of *mother*, because a certain ideology of gender roles and familial relations is prevalent in our society (for better or worse - this concept of ideology is descriptive, not pejorative) and one of its elements is the assumption that parenting is and should be gendered. *Welfare* invokes stereotypes about Black Americans, and it is a crucial aspect of the ideological conditions of American politics that greater salience accrues to stereotypes such as “Blacks are lazy” than, say, “Black culture is a culture of resilience in the face of systemic injustice”.

\textsuperscript{23} Anderson speaks of representations rather than propositions and distinguishes various forms of public standing, but these details are not relevant to my purposes.
The complex relations between language, inference, context and ideology cannot be discussed here even fragmentarily (not only for reasons of space, but because there are still many open questions in this area), but the most relevant aspect is that ideologies are operative even if they are not endorsed. As Swanson puts it, “relationships between language and ideology [...] are often not part of language users’ shared information, and often not part of any intentions they could articulate or even uncover.” None of this is to say that problematic or opaque social meanings cannot be brought to awareness, disavowed, or challenged. But there is a price: at the very least that of disrupting the smooth flow of a conversation through making utterances that deliberately aim to counter the significance awarded to them by the ideologically-shaped discursive context.

Although I have employed inferentialism primarily in a technical way as a convenient tool to describe the interaction between social meanings and conversational context, herein lies the value of this framework on a more philosophical level. On the inferentialist picture, to understand an utterance is, at least partly, to understand what (material) inferences it licenses. The latter is importantly co-constituted by salient elements of the discourse context. The background assumptions, stereotypes, prejudices etc. to which social meanings are tied are not merely evoked by certain linguistic expressions as a by-product of communicative acts; they are in an important way part of the meaning of those acts—if by meaning we understand not just literal content plus various pragmatic enrichments, but the total impact of an utterance on the discourse. Moreover, once the background premises are raised in salience, they remain so, at least for some stretch of discourse afterwards, co-determining the meaning of further utterances, including those by other interlocutors or those who would disagree with the first speaker. 24 Inferentialism thus construed is a strictly externalist framework, positing that we do not have full control over, or full insight into what our words mean—although it is what could be called an “ideological externalism”, not necessarily equivalent to the more familiar referential variety. A fuller development of this externalist understanding of ideology in language is left for future work.

Acknowledgements I received many helpful comments on earlier drafts of this paper. Special thanks go to Katarzyna Kijania-Placek, Krzysztof Posłajko, Eric Swanson, the audiences at the Value in Language workshop, PhiLang 2021 conference and the Jagiellonian University Institute of Philosophy, and the two anonymous referees at Synthese. Work on this project was supported financially by the National Science Centre, Poland under grant 2019/32/C/HS1/00475.

Declarations

Conflict of interest There is no conflict of interest.

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

24 On the difficulty of lowering the salience of things made salient—of unringing the metaphorical bell—see Simpson (2013).
by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Anderson, E. (2010). *The imperative of integration*. Princeton: Princeton University Press.
Anderson, L. (2018). Calling, addressing, and appropriation. In D. Sosa (Ed.), *Bad Words*. Oxford: Oxford University Press.
Anderson, L., & Lepore, E. (2013). Slurring words. *Noûs*, 47(1), 25–48.
Bach, K. (2018). Loaded words: On the semantics and pragmatics of slurs. In D. Sosa (Ed.), *Bad Words*. Oxford: Oxford University Press.
Beltrama, A. (2020). Social meaning in semantics and pragmatics. *Language and Linguistics Compass, 14*, 9.
Bolinger, R. J. (2017). The pragmatics of slurs. *Nouss*, 51(3), 439–462.
Brandom, P. (2000). *Articulating reasons: An introduction to inferentialism*. Cambridge, MA: Harvard University Press.
Brandom, R. (1994). *Making it explicit: Reasoning, representing, and discursive commitment*. Cambridge, MA: Harvard University Press.
Cepollaro, B., & Stojanovic, I. (2016). Hybrid evaluatives: In defense of a presuppositional account. *Grazer Philosophische Studien, 93*(3), 458–488.
Cepollaro, B., & Thommen, T. (2019). What’s wrong with truth-conditional accounts of slurs. *Linguistics and Philosophy, 42*, 333–347.
Cepollaro, B., & Torrengo, G. (2018). The worst and the best of propaganda. *Disputatio, 1*(51), 289–303.
Davis, C., & McCready, E. (2020). The instability of slurs. *Grazer Philosophische Studien, 97*(1), 63–85.
Diaz-Leon, E. (2020). Pejorative terms and the semantic strategy. *Acta Analytica, 35*, 1.
Drobnák, M. (2020). Inferentialism on meaning, content, and context. *Acta Analytica, 35*(1), 35–50.
Dummett, M. (1973). *Frege: Philosophy of language*. London: Duckworth.
Eckert, P. (2018). *Meaning and linguistic variation: The third wave in sociolinguistics*. Cambridge University Press.
Fraser, N., & Gordon, L. (1994). “Dependency” demystified: Inscriptions of power in a keyword of the welfare state. *Social Politics: International Studies in Gender, State & Society*, 1(1), 4–31.
Hall-Lew, L., Moore, E., & Podesva, R. (2021). Social meaning and linguistic variation: Theorizing the third wave. Cambridge: Cambridge University Press.
Haslanger, S. (2012). *Resisting reality, social construction and social critique*. Oxford: Oxford University Press.
Haslanger, S. (2014). Social meaning and philosophical method. *Proceedings and Addresses of the American Philosophical Association, 88*, 16–37.
Henderson, R., & McCready, E. (2018). How dogwhistles work. In *New Frontiers in Artificial Intelligence*, pages 231–240.
Hess, L. (2020). Practices of slur use. *Grazer Philosophische Studien, 97*, 86–105.
Hess, L. (2021). Slurs: Semantic and pragmatic theories of meaning. In P. Stalmaszczyk (Ed.), *The Cambridge handbook of the philosophy of language*. Cambridge: Cambridge University Press.
Hom, C. (2008). The semantics of racial epithets. *Journal of Philosophy, 105*(8), 416–440.
Hom, C. (2012). A puzzle about pejoratives. *Philosophical Studies, 159*(3), 383–405.
Hornby, J. (2001). Meaning and uselessness: How to think about derogatory words. *Midwest Studies In Philosophy, 25*(1), 128–141.
Jeshion, R. (2013). Slurs and stereotypes. *Analytic Philosophy, 54*(3), 314–329.
Kahan, D. M. (1997). Social influence, social meaning, and deterrence. *Virginia Law Review, 83*(2), 349–395.
Kaluziński, B. (2020). Inferentialism, Context-Shifting and Background Assumptions. *Erkenntnis*.
McCready, E. (2010). Varieties of conventional implicature. *Semantics and Pragmatics, 3*(8), 1–57.
Murray, S. E. (2014). Varieties of update. *Semantics and Pragmatics, 7*(2), 1–53.
Nunberg, G. (2018). The social life of slurs. In D. Fogal, D. Harris, & M. Moss (Eds.), *New work on speech acts*. Oxford: Oxford University Press.
Popa-Wyatt, M. (2020). *Slurs, pejoratives, and hate speech*. Oxford Bibliographies.
Potts, C. (2005). *The logic of conventional implicatures*. Oxford studies in theoretical linguisticsOxford: Oxford University Press.

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

Rappaport, J. (2020). Slurs and toxicity: It’s not about meaning. *Grazer Philosophische Studien, 97*(1), 177–202.

Saul, J. (2018). Dogwhistles, political manipulation, and philosophy of language. In D. Fogal, D. Harris, & M. Moss (Eds.), *New work on speech acts*. Oxford: Oxford University Press.

Sennet, A., & Copp, D. (2015). What kind of a mistake is it to use a slur? *Philosophical Studies, 172*(4), 1079–1104.

Simpson, R. M. (2013). Un-ringing the bell: Mcgowan on oppressive speech and the asymmetric pliability of conversations. *Australasian Journal of Philosophy, 91*(3), 555–575.

Stanley, J. (2015). *How propaganda works*. Princeton: Princeton University Press.

Stanley, J. (2018). Replies to Cepollaro and Torrengo, Táiwò, and Amoretti. *Disputatio, 10*(51), 345–359.

Swanson, E. (ms.). Slurs and ideologies (forthcoming). In Celikates, R., Haslanger, S., and Stanley, J., editors, *Analyzing Ideology*. Oxford University Press, Oxford.

Tirrell, L. (1999). Derogatory terms: Racism, sexism and the inferential role theory of meaning. In K. Oliver & C. Hendricks (Eds.), *Language and liberation: Feminism*. Philosophy and Language: SUNY Press.

Tirrell, L. (2017). Toxic speech: Toward an epidemiology of discursive harm. *Philosophical Topics, 45*(2), 139–162.

Valtonen, P. (2019). Generic inferential rules for slurs: Dummett and Williamson on ethnic pejoratives. *Synthese*.

Whiting, D. (2008). Conservatives and racists: Inferential role semantics and pejoratives. *Philosophia, 36*(3), 375–388.

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

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

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