CrossMark

Slurs, roles and power

Mihaela Popa-Wyatt[1] (iD) · **Jeremy L. Wyatt**[2]

Abstract Slurring is a kind of hate speech that has various effects. Notable among these is variable offence. Slurs vary in offence across words, uses, and the reactions of audience members. Patterns of offence aren't adequately explained by current theories. We propose an explanation based on the unjust power imbalance that a slur seeks to achieve. Our starting observation is that in discourse participants take on *discourse roles*. These are typically inherited from social roles, but only exist during a discourse. A slurring act is a speech-act that alters the discourse roles of the target and speaker. By assigning discourse roles the speaker unjustly changes the power balance in the dialogue. This has a variety of effects on the target and audience. We show how these notions explain all three types of offence variation. We also briefly sketch how a role and power theory can help explain silencing and appropriation. Explanatory power lies in the fact that offence is correlated with the perceived unjustness of the power imbalance created.

✉ Mihaela Popa-Wyatt
   mihaela.popawyatt@gmail.com

   Jeremy L. Wyatt
   jeremy.l.wyatt@gmail.com

[1] Department of Philosophy, University of Birmingham, Birmingham B15 2TT, UK

[2] School of Computer Science, University of Birmingham, Birmingham B15 2TT, UK

1 Introduction

Slurring is a kind of hate speech that has special properties. There are thousands of terms that are identified as slurs,[1] including those based on race ("nigger", "chink"), gender ("bitch"), nationality ("limey", "boche"), and sexuality ("faggot", "dyke")". Slurs can harm and degrade their targets, making them feel humiliated, dehumanised, disempowered, and silenced. Slurs may also offend non-targets, often making them feel complicit, while meeting with approval from bigots. A slur can also be reclaimed by the target group, so as to express solidarity and pride. Slurs are interesting precisely because of this unusually rich set of effects.

Slurs vary in the strength of these effects. In particular, there is considerable variation in the degree of offence caused. Offence varies across *different slur words*, across *different uses of the same slur word,* and across the reactions of *different audience members*. Current theories struggle to parsimoniously explain the resulting patterns of offence. The contribution of this paper is to remedy this. A secondary contribution is to show that the same mechanism also explains silencing and appropriation.

The core of the explanation is that many slurring utterances are not just about expressing contempt, but are also oppressive speech.[2] The offence caused by a slurring utterance varies with the degree of oppression that it plausibly seeks to achieve. Slurs vary in their ability to be so used, from the most potent, such as "nigger" and "faggot", to the milder, such as "limey" and "yankee".[3] All are identified as slurs, and all can be used to convey contempt, which causes a baseline offence. Additional offence is caused by those slurs used to oppress. We refer to any slur that can plausibly be used to cause some degree of oppression as an *oppressive slur*.

Oppression can be explained as unjust power using the *power-over* conception of power (Weber 1978; Dahl 1957). According to this theory, power is the extent to which a person A can affect a person B contrary to B's interests (Lukes 1974). Thus, A's power over B becomes oppressive if it is used to cause harm to B. The goal of oppressive slurring is to achieve and maintain unjust power over the target. A slurring use draws and builds on historical and/or contemporary oppression of the target group.[4]

---

[1] The racial slur database contains more than 2500 racial slur terms, including derivatives. Jeshion (forthcoming) estimates that 560 of these refer to people of African descent, 53 to those of Chinese descent, and that higher numbers indicate stronger community backing and impact on the target group.

[2] Our account relates to accounts of oppressive speech and discursive injustice by Matsuda et al. (1993), Hornsby (1995), Tirrell (1999, 2012), McGowan (2009, 2012), Maitra (2012), Kukla (2014), Langton (2012), and accounts relating slurs to ideology (Stanley 2015; Swanson ms.), and to analyses of subordination and silencing of women through pornography (Langton 1993; Maitra 2009, among others).

[3] Clearly, it is the case that there is a significant grey area, the transition from non-oppressive to oppressive slurs is not sharp.

[4] The oppression drawn upon is structural—(Young 1990; Cudd 2006; Haslanger 2012; Anderson 2012). Once oppression has become institutionalised through laws, customs, and practices, unjust harm can occur, whether or not the individuals maintaining those practices have oppressive intentions.

Most authors put contempt expression and derogation at the heart of what slurs are used for.[5] Slurs are used for this, but contempt alone can't explain the greater offence, nor the oppressive effects, of some slurring acts. Adding *unjust power* explains phenomena that contempt can't. Power is not, however, enough on its own. Slurring uses are rich enough in their effects that a single mechanism cannot account for all their properties. We should seek a set of relatively simple, interlinked mechanisms, that work together to explain the various effects of slurring utterances.

We will argue that oppressive slurs achieve oppression via an act of *role assignment*.[6] The speaker takes the dominant role, while the target is assigned the subordinate role. The notion of role assignment is best explained using speech-act theory. Much of the paper shall be concerned with setting up the machinery we need to support this. We suggest that the speech-act is best thought of as an exercitive assignment of a *discourse role*. Discourse roles determine the status and rules that govern the dialogue behaviour of participants. Thus, an exercitive assignment of a discourse role is a meta-move that changes the rules governing other dialogue moves.

To see the need for this approach, we must first understand variable offence and how current theories fare regarding it (Sect. 2). We shall then detail our theory of slurs as oppressive speech acts that alter power via discourse role assignment (Sects. 3–6), and show how this theory's predictions fit the data on variable offence, silencing, and appropriation (Sects. 7–8).

2 Variable offence

We must first mention a key assumption. In some works (e.g. Hom 2008, 2012) the focus in describing variability is on the derogative content of the word. Hom distinguishes derogation and offence as follows: "offence is a psychological result on the part of the discourse participants, and is a function of their beliefs and values … derogation is an objective feature of the semantic contents of pejorative terms" (Hom 2012: 397). We would suggest that, to avoid committing too early to a semantic or pragmatic solution, we should relax Hom's statement, and think about derogation simply in terms of it being part of speaker meaning. So, derogation is part of the speaker's communicative intentions, while offence is an achieved effect on the audience members, determined in part by their beliefs and values.

As noted, slurs vary in the degree of offence they cause. As Jeshion (2013b: 309) says, "'nigger' is said to be more offensive than 'chink', as well as 'spook' and 'jigaboo', terms used for the same socially relevant group. 'Kike', 'yid', and 'hymie' are said to differ in their offensive intensity." Thus, offence doesn't just vary across different slur words for different groups (*inter-group variation*), but also across different slur words for the same group (*intra-group variation*). We refer to

---

[5] Hornsby (2001), Saka (2007), Hom (2008), Richard (2008), Jeshion (2013b), Camp (2013), among others.

[6] Note that we are not claiming that words are only slurs if they create an unjust power imbalance. Slurs include words that are not necessarily oppressive.

these collectively as *variability type-1* (VT1) or *word-variation*. This is the type of variable derogation which previous authors have tried to explain.

Offence doesn't just vary across different slur words. It also varies across *different uses* of the same slur word, depending on speaker and context. We refer to this as *variability type-2* (VT2) or *use-variation*. Consider the following examples (*C* marks contemptuous intonation, and *NC* non-contemptuous, friendly tone of voice):

(1)   White person to African-American person: "You are one useless nigger^C."
(2)   White person to African-American person: "Yo, my nigger ^{NC}, missed you."
(3)   African-American person to African-American person: "You are one useless nigger^C."
(4)   African-American person to African-American person: "Yo, my nigger^{NC}, missed you."

What offence would we expect these utterances to cause? In considering this we must, since offence is a psychological property of the hearer, carefully define the meaning of statements such as "slur X is more offensive than slur Y". Since offence is a psychological and thus subjective property, such statements can only be meaningfully defined with respect to a particular individual. However, by taking idealised individuals, e.g. a non-racist, non-homophobic, non-sexist individual, we can make statements about the ranking of offensiveness of slurring utterances taken by that person. Thus, we can make statements such as 'slurring utterance X is more offensive than utterance Y to an idealised person A, where A is a non-racist, non-homophobic, non-sexist individual'. This is how we define statements about the relative offence of particular slurs in this paper. However, as a short-hand we will simply write statements of the form "utterance/use X is more offensive than utterance/use Y", and the reader should simply fill in the caveat regarding the ranking of subjective offence taken by the idealised non-bigoted individual defined above.

Having taken this into account, let us consider the utterances (1)–(4). First, consider contemptuous versus non-contemptuous uses. The contemptuous uses (1) and (3) are more offensive than the non-contemptuous uses (2) and (4). Now consider the in-group versus out-group uses. The out-group contemptuous use (1) is widely considered more offensive than the identical in-group use (3), and the same is true of the corresponding pair of non-contemptuous uses (2) and (4).

The issue is to explain why, if the speakers of (2) and (4) don't intend to convey contempt, there is a difference in the offence we would expect. The same can be said for (1) and (3), where contempt is intended, but where (1) is clearly more offensive. We can make many variations to the context in which these utterances are made and the relative offence within each pair will remain the same.[7] For example, suppose that the speaker and target of utterances (2) and (4) are close friends, or that alternatively they are strangers. In either case (2) will be more offensive than (4). Now suppose we adjust (1) and (3) so that they are not uttered with a contemptuous

---

[7] Other contextual factors might be: speaker's accent, familiarity between speaker and hearer, and whether they belong to a close-knit social or cultural group. We are not making an empirical claim, but merely delineating dimensions of use-variability.

tone, (1) is still more offensive than (3). This shows that the degree of contempt or derogation conveyed by a slur word cannot be the only ingredient in variable offence. Pragmatic factors, such as the *speaker's group-membership*, are also important.[8]

Finally, there is a third type of variability, not previously identified in the literature. This is that the same slurring utterance may cause different levels of offence to different audience members. Imagine a slurring utterance in front of an audience who vary from the highly bigoted to non-bigoted. It is reasonable to assume that different audience members will be offended to different degrees. We refer to this offence variation as *variability type-3* (VT3) or *audience-variation*.

3 Variable offence and current theories

How shall we explain VT1, VT2, and VT3? There are numerous accounts of slurs. These are mostly concerned with whether slurs' derogatory power is a matter of meaning, and on which side of the semantics/pragmatics divide to locate that meaning.[9] As variable offence is our focus, we will review accounts that directly address this issue. Our strategy is to argue that none has sufficient resources alone to adequately explain all three varieties of variable offence. We offer a theory that provides sufficient pragmatic resources to fill the explanatory gap required.

3.1 The truth-conditional semantic approach

Starting with semantic theories, we take Hom (2008) as representative. Take VT1, there are two types of variation falling within VT1: intra-group and inter-group. Inter-group variation is well explained by Hom. Hom talks not directly of variable offence, but of variable derogation. On that account, slurs express that the target "ought to be subject to [deontic propositions about discrimination] $p_1^* + \cdots + p_n^*$ because of being $d_1^* + \cdots + d_n^*$" all because of being part of such-and-such-group" (2008: 431). As Hom notes, "the $d_1^* + \cdots + d_n^*$ are the negative properties derived from the racist ideology." These form a negative stereotype of the target group. Hom also makes clear that "the derogatory force … is directly proportional to the content of the property it expresses, which is in turn directly proportional to the turpitude and scope of the supporting racist institution" (2008: 432). While Hom and May (2013: 10) later argue that it is not the meaning itself that is offensive— offence being a psychological property—it is still clear that they view

---

[8] These offence variations should be experimentally quantified.

[9] Aside from Anderson and Lepore (2013a), most authors provide an account of the meaning of slur words. Some treat meaning as propositional content—whether as part of *truth-conditional semantics* (Hom 2008) or *non-truth-conditional semantics* (Saka 2007). Other treatments are as *editorial comment* (Bach forthcoming), *presupposition* (Schlenker 2007), and *conventional implicature* (Williamson 2009; Whiting 2013; Lycan 2015). Others take the meaning to be *non-propositional* and *non-truth-conditional*, differing on the fundamental ingredient: e.g. *contempt expression* (Barker 2000; Copp 2001; Richard 2008; Jeshion 2013b), *derogatory perspectives* (Camp 2013), *register* (Predelli 2013), *expressive meaning* (Potts 2007), *identity-affiliation* (Nunberg forthcoming), *contrastive-choice* (Bolinger 2015).

"offensiveness as a property of utterances" and admit that "word meaning could be part of an explanation of offensive uses of these words". On this ground, it is reasonable to suggest that word meaning, insofar as it does explain inter-group offence variation, relies on an increase in derogatory force to cause greater offence.

Thus, in simplified terms, offence can vary with word meaning because of semantic encoding of more or less negative stereotypes for different groups. The pejorative content is a function of the stereotype and the associated racist beliefs and ideology about how the target ought therefore to be treated. Thus, "nigger" and "chink" are offensive to different degrees because the racist stereotypes predicated are more negative for one group than the other. Note that this means that the differences in offence ultimately rely on different institutions of racism: "variation in derogatory force associated with different epithets is a result of having different racist institutions causally support their predicative content" (Hom 2008: 433).

Because of this reliance, Hom's explanation cannot account so easily for variation across different slurs for the same group (intra-group variation). For example, "nigger" is considered more offensive than "spook". But for Hom, in order for it to be so, it must be the case that the stereotypes encoded by the slur rest on two different racist ideologies. This is implausible, because it means that there must be a different institution of racism for "nigger" to that for "spook", as argued by Anderson and Lepore (2013a: 361). So Hom has mixed success in explaining VT1.

VT2 is a bigger limitation for semantic theories. This is because if the stereotype is encoded semantically, and as far as semantics is able to explain offence, then the offence caused should be the same, whoever says it. This isn't the case. Take the examples (1)–(4). We saw that variations in speaker caused offence to vary. Then how is the semantic theorist to explain VT2?

Hom appeals to the idea that slurs can have more than one meaning. So, for example, to enable appropriated in-group uses they may acquire an additional, non-derogatory meaning. For example, consider the word "queer". It has two meanings: one derogatory, one not. However, this fails to explain two things. First, why is it the case that only in-group members can access one of these meanings? Having meanings that only some language users can access seems odd. Second, what happens during the process of appropriation? During this process bigots still used the term "queer" contemptuously, and this contemptuous use can continue even once reclamation is complete. But if derogation resides in semantic meaning, and not pragmatic effects, then how can Hom explain that "queer" can be used both derogatorily and non-derogatorily during the reclamation process? Presumably during reclamation, "queer" has not yet acquired a non-derogatory meaning, and thus the word has only one, derogatory, meaning.

To explain this, Hom could argue that for in-group uses—prior to reclaim—the non-derogatory meaning is recovered via conversational implicature. The reasoning might go something like this: "the speaker has just called me a queer, but the speaker is also gay, and he is not speaking with a contemptuous tone, so perhaps he is applying it in some positive way." However, conversational implicature falls down in explaining why a friendly out-group speaker still causes offence. The target would correctly infer that the friendly tone means that the out-group speaker didn't

intend to derogate them. But they are still offended, so the group-membership of the speaker has trumped their intention. Thus, the semantic account can't be saved by conversational implicature.

What about VT3? Offence is at least partly caused by the stereotype, and this varies with the racist institution. Thus, Hom can argue that an audience member's offence will vary inversely with their degree of subscription to this racist institution. However, if the stereotype is semantically encoded, we should expect that competent users know its precise details as part of their semantic understanding. This is implausible: lots of users are equally offended without having a common, precisely similar stereotype in their semantic understanding (Jeshion 2013a). However, following Burge (1986), Hom could deny this, holding that competent users can be partially ignorant or mistaken about the precise nature of these stereotypes and norms. Thus, listeners can be offended while only knowing partial details of the stereotype, or the detailed history that underpins it. As long as the hearer's partial understanding is good enough for them to understand sufficiently well some of the difference between the extent to which different words encode a subscription to a racist stereotype, then VT3 can be accounted for semantically.

In at least two cases not all of the offence variation can be explained by semantics alone. If we want to follow Hom, we must at least admit that there are gaps that semantics can't fill alone. These must be filled pragmatically.

3.2 The expressivist approach

An alternative to a truth-conditional semantics is an expressivist semantics. Here, we take Jeshion's (2013a, b, forthcoming) account as representative. In expressivist theories the main function of a slur is to *conventionally express contempt* for the target. For Jeshion, this contempt expression is part of semantics.[10] Critically, for the topic of this paper, she maintains that this semantic expression of contempt is uniform across all slurs, in that a common core attitude is expressed and there is no semantic "intensity marker" attached to a slur to indicate variation in contempt. Thus, semantics cannot explain variable offence as being caused by variable contempt. Critics (e.g. Hom 2008) claim that this makes it impossible for expressivism to explain variable offence, and thus reject it.

There are two possible responses to this. First, is it possible to have an expressivist semantics that allows contempt expression to vary across terms? On such an account, contempt would have a semantic intensity marker, and so "nigger" would be more offensive than "spook" because it expresses a greater degree of contempt.[11] This is tempting, but there is a problem. If this were the case, we would expect the degree of offence to follow from the degree of contempt a word is able to convey. But it does not. For example, there are many words that slur the rich and powerful, such as "toff", "Lord Snooty", "nob", "suit", "the one percent" and

---

[10] This does not mean that the speaker is *asserting* that she has contempt for the target, on account of their being part of the target group.

[11] This might be implemented by applying Potts' (2007) model of "expressive indices"; see Popa-Wyatt (2016).

"Hooray Henry", that all convey considerable contempt. Yet these are socially acceptable. Degree of contempt cannot therefore be the right answer. We have to explain why words that convey equal contempt are not equally offensive.

As Jeshion correctly points out, the answer lies in pragmatics (2013b). She suggests the Derogatory Variation-Utterance thesis: "utterances of different slurring terms engender different degrees of intensity of offensiveness" (2013b: 320). She then suggests several different pragmatic (*perlocutionary*) effects that could account for derogatory variation. These are: the pragmatic (rather than semantic) activation of stereotypes (2013b: 321), the compounding effect of a slur on the history of oppression (Saka 2007: 148; Jeshion 2013b: 322), and the variable offence caused by breaking prohibitions of varying strength (2013b: 323). Thus, different pragmatic resources could be used to account for VT1 (e.g. history of oppression evoked by the word and group stereotypes), VT2 (e.g. varying levels of taboo for different uses), and VT3 (e.g. individual variations across audience members of the pragmatics underpinning VT1 and VT2). This is reasonable, but then there isn't just one pragmatic mechanism that covers all the relevant dimensions of variability; rather, distinct mechanisms account for each kind of variability. We should prefer a more compact explanation.

3.3 The non-content approach

So far we have dealt with content theories. Do non-content theories do better at covering all cases of variable offence? The leading contender is prohibitionism (Anderson and Lepore 2013a, b). This starts with the observation that slurs are still offensive when they occur in belief and speech reports, such as:

(5)   *Eric said that a bitch ran for President of the United States in 2008.*

Anderson and Lepore claim that this doesn't automatically capture the offence of Eric's utterance, but that the reporter will certainly offend. They argue that since content-based accounts predict that slurs should not offend when so embedded, and that since they do, then content theories must be wrong.[12] If the intuition about the data's pattern of offence is correct, this is compelling. Prohibitionism is their alternative. It says that slur words are offensive simply because they are prohibited (taboo).

How well does prohibitionism match variable offence phenomena? It tries to accommodate VT1 by saying that there is a correlation between offence variation across slur words and variation in their degree of taboo. It tries to accommodate VT2 by saying that although slurs are taboo, the taboo can be suspended depending on the context or speaker. So, for example, in-group uses come with a "defeasible escape clause" (2013a: 42). Finally, although Anderson and Lepore do consider this, VT3 could also be accounted for by saying that different audience members have different levels of individual taboo. Prohibitionism is flexible, many additional

---

[12] They make other arguments, which we have no space to describe, but this is a sufficient example to see their general strategy.

clauses about taboo suspension or reduction can be created to fit the data. Thus, we could create rules that non-contemptuous uses are less taboo than weapon uses; in-group uses are taboo-free; and mentioning, rather than using, may (or may not) be taboo, depending on who mentions the slur, their company, and the context (e.g. legal, academic, political).

So, if taboo can be made to fit the data what's wrong with this explanation? Taboo lacks explanatory power. The advert for prohibitionism is its simplicity. This is deceptive. Consider all the various defeasible escape clauses that are required to rank offensiveness across our three types of variability. In examples (1)–(4) we saw that in-group uses will require an escape clause. So, even if a white and black person are close friends, use of "nigger" may still offend. But consider further cases. It may not offend, even if they are not friends, if the target permits its use. Furthermore, African-Americans who are only passingly familiar may use the term to refer to each other without offence. But, in another case, reported in the press (Gabbatt 2013), an African-American woman was offended, and successfully sued her African-American manager for calling her "nigger". A complex web of exceptions and non-exceptions must be written to cover all these cases. Prohibitionism is like Ptolemy's earth-centric model of the solar system: the basic idea is simple, but to fit the data the exception clauses must rise in number like epicycles. Prohibitionism isn't parsimonious. It merely echoes the facts of variable offence rather than explaining them compactly. We should like a compact explanation of why words are differentially taboo.

Another way of saying this is to pose the question: if offence variability is explained by the degree of prohibition, what explains that degree of prohibition? Anderson and Lepore try to get round precisely this objection, that "prohibitionism gets it backwards; it is the offence that explains the prohibition, not the other way round" (2013b: 354). They counter by pointing out that there are examples where prohibition precedes offence, for example blasphemy. But, as they recognise, this fact alone doesn't compel, we must establish why this is the case for slurs. As a possible answer, they say "it might be that groups prohibit names not explicitly adopted by them, for calling a group by a name that its members have not chosen may be viewed as an attempt to usurp their authority to choose" (2013b: 355). This is not compelling. If this were the case surely all slurs would be equally offensive, since all would violate the same prohibition. Their final improvement is to note the effect of subordination: "in situations where one group is in a subordinate position to another, uses of an expression by the dominant group to refer to the subordinate group can provoke offence" (2013b: 355). This is insightful, but it puts the crack in the dam. The difficulty is that once we have admitted, as we must, that the prohibition arises from subordination we are better to turn to subordination as the core explanation.

Prohibitionism is unsatisfying in another way. It is silent on why slurs have different effects on audience members and targets than other pejoratives. This leaves it impoverished. What is interesting about slurs is all the other effects they have. They silence, dehumanise, assault, and oppress. On these effects prohibitionism is mute. Thus, it might be a part of a hybrid theory, but it cannot be the core. A satisfactory theory of slurs must explain these other properties. This leads us to

our central question: does a parsimonious, pragmatic account of offence variation exist? The answer to this question is yes. It is based on a speech-act account of unjust power imbalance. We will set it out, and then show how it explains the data.

4 Roles and power

We contend that slurring utterances seek to create (or maintain) an *unjust power imbalance* via role assignment. Our second contention is that the degree of offence caused is correlated with the magnitude of the perceived unjustness of the power imbalance associated with this role. So, what's in a role?

Roles are social constructs that carry information about permissible and expected behaviours, social status (i.e. rank relative to other roles), rights, and responsibilities.[13] As Goffman notes, "role" is the basic unit of socialisation: "it is through roles that tasks in society are allocated and arrangements made to enforce their performance" (1961: 77). In society we allocate, by various means, many social roles to individuals. Human interactions are informed by the expectations these create. Roles also often come in pairs or groups, and convey information about relative status and power. When roles change, power can change. While in many situations roles with power imbalance are accepted and useful (e.g. parent and child), others have unjust power imbalances that are created and maintained by force (e.g. master and slave).

Different social roles taken on by an individual persist over different time periods. Some commonly last an entire lifetime (ethnicity, gender), others typically last for many years (profession, being a spouse), while some last for months (pregnancy, trainee), days (being a workshop attendee, tourist, jury member) or hours (being a customer in a restaurant). We can thus refer to the duration of a social role, and thus to—relatively speaking—long-term social roles and short-term social roles. During the period within which a person holds a role some number of individual social interactions will occur. A discourse is a type of individual social interaction, and so a role that lasts for a single discourse is thus a very short-term social role.

There is an intimate relationship between social roles and discourse. Social roles structure social interactions, and are structured by them. Discourse is structured by social roles in several ways. Speakers signal information about the status of a social role via honorifics and etiquette terms.[14] In interpreting what speakers say, hearers are guided by information about their interlocutors' social roles. A socially dominant speaker can rule out communicative possibilities for his dominated interlocutors (Lewis 1979: 347). Finally, speakers can also say things that alter social roles. Examples include declaring two people to be married, convicting a criminal, and investing someone into a particular office, such as a monarch, an

---

[13] Social roles are stable constructs, whether or not they are institutionalized (Mallon 2003).

[14] Honorifics serve to establish power in discourse, e.g. when a speaker addresses someone with "Sir", this signals to peers that they should treat that person with deference. Potts (2007) provides a semantic modelling.

admiral, or a priest. Such role assignments are performatives in that they allocate a long-term social role to a person, i.e. the role persists beyond the temporal duration of the discourse in which it is allocated.

It is necessary to distinguish the social roles that speakers draw on from their use in discourse. To do this, we introduce the notion of *discourse roles*. We will first sketch the concept and then show its utility. Since discourse is a social interaction, discourse roles are precisely short-term social roles that exist only for the purposes of a particular dialogue. They often inherit from longer-term social roles which participants hold externally to the conversation. Thus, a person may have many social roles external to a particular conversation (employee, policeman, homicide detective, parent, man, African-American). Which of these external roles is salient within a particular conversation depends on the circumstance. So, someone talking with their boss adopts a different external social role (employee) as their discourse role than when talking with their child (parent). Discourse roles may be drawn from social roles that neither of the participants actually have. For example, discourse roles are sometimes play-acted, such as when an adult play-acts as a child, thus inviting their interlocutor to take on a complementary role (e.g. a parent, or another child). Finally, discourse roles can also be changed during a conversation.

A discourse role has three parts: *discourse status*, *discourse rules*, and the associated *social role*. Discourse status denotes relative rank between a role and other roles. It reflects the fact that human society is hierarchically organised, and that many social interactions involve participants with a clear difference in rank. Discourse rules govern expected moves, permissible moves, their expected effects, and rules of interpretation. The same utterance—in an otherwise identical situation—will be interpreted differently depending on the speaker's discourse role. For example, a person saying to a bank teller "I want $1000" will engender a different interpretation depending on whether they are a customer or a bank robber. A speech-act that changes the discourse role of a participant is thus important because it changes the discourse rules and discourse status for that participant. It also brings to salience the associated long-term social role.

What work do discourse roles actually do? Why is it necessary to separate long-term social roles external to the dialogue from the short-term social roles (discourse roles) that are internal to the dialogue? The answer is that without such a distinction we cannot have a coherent theory of how roles determine dialogue. Let's take an example from the film "In the Heat of the Night". After being arrested, in Mississippi, in the 1960s, on suspicion of murder, the Sidney Poitier character, Mr Tibbs, is taken to the police station. Mr Tibbs has four external social roles that are relevant: first, he is an African-American in the deep south during a period of overt racist practice; second, he is briefly[15] a murder suspect; third, he is a police officer; and fourth, he is the number one homicide expert with the Philadelphia police department. The conversation with the chief of police begins with Mr Tibbs

---

[15] Briefly, but still prior to the start of the dialogue in question, hence this is a 'long-term' social role for Tibbs that exists external to the conversation.

as a murder suspect, and the fact of his race determining the conversational trajectory, and the balance of power:

> PC: Got a name boy?
> MT: Virgil.
> PC: Virgil? <laughs> I don't think we're going to have any trouble, are we Virgil?
> MT: No.
> PC: What d'you hit him with boy?

It should be clear that the power lies with the police chief. As the conversation progresses it transpires that Mr Tibbs is a police officer:

> PC: Just what you do up there in Pennsylvania, to earn that kind of money?
> MT: I'm a police officer. <shows police badge>
> PC: Oh. Yeah. <scratches head, sits down>

At this point the tone of the conversation changes. The fact that he is a policeman and a homicide expert now determine the trajectory of the discourse. This is shown with particular force, when, after further conversation, the police chief asks him to look at the body of the murdered man as a favour:

> PC: Look, if they pay you $162.39 a week to look at bodies. Why can't you look at this one?
> MT: Why can't you look at it yourself?
> PC: Because I'm not an expert, officer.

Although the tone of the conversation has already changed, "officer" is the first time Virgil Tibbs is referred to with an honorific. It is a clear statement of respect. The request for a favour demonstrates a significant shift in the power balance during the conversation.

This example shows that the salience of external social roles changes during a discourse. To make sense of the conversation at any moment, we have to know which external social roles of Mr. Tibbs are salient at that moment. Certain external roles became less salient (murder suspect, member of an oppressed racial group), while others (police officer, homicide expert) became more salient. But three of the actual external social roles of Mr. Tibbs did not change: all the way through the conversation, and before and after, he was an African-American, a police officer, and a homicide expert.[16] Thus, to describe the conversational kinematics we must separate a participant's external social roles, that exist beyond and throughout the dialogue, from their corresponding changes in salience during the dialogue, and the way that those changes in salience cause the conversational kinematics to change.

On the model proposed here, a discourse role is created when an external social role becomes salient, and extinguished when the discourse ends, or when the external social role ceases to be salient. We must, however, separate the discourse and external social roles themselves. This is because their assignments last for

---

[16] Only in the conversation did his external role as a murder suspect change.

different durations. When a discourse role is assigned it can only last as long as the discourse. When an external social role is assigned it can last beyond the discourse. If we only had external social roles we would not be able to explain the difference in effect between the following two utterances, each of which makes a role assignment:

(6)   *We find you guilty of the charge of murder.*
(7)   *You dirty nigger.*

In (6) the speaker (a jury member) assigns the roles of criminal and convicted murderer to the target. In (7) the speaker (a work colleague) assigns a subordinate role to the target by slurring. The effect of (6) lasts the rest of the life of the target (unless the conviction is quashed on appeal). The effect of (7) does not. If we were to only have external social roles, then there is no criterion to distinguish why one assignment lasts and one doesn't. Thus, if we are to have the notion of role assignment in discourse, discourse roles are necessarily different from (though closely related to) the external social roles from which they inherit.[17]

It is critical, however, to understand that these discourse roles are indeed short-term social roles that exist internally in the discourse. Hence we use the term 'discourse role' as a short-hand for this short-term social role that is internal to a specific dialogue. The discourse role is a social role that comes into being when it is made salient in a dialogue, and which does not last beyond the dialogue. A discourse role may, of course, result in the creation of an external social role that can last a long time. This is exactly what happens when, for example, a couple are married. In the moment that the discourse roles husband and wife are created, a pair of long term-social roles, that will persist after the ceremonial conversation has ended, are also created. We will hereafter use the terms discourse role and external social role or simply social role.[18]

The conversational kinematics at each moment are determined by the combination of discourse roles currently held by each participant. In the first excerpt Mr Tibbs is expected, under his discourse roles as a murder suspect and a black man in Mississippi, to be subordinate and to respond with respect to his white counterparts,

---

[17] Discourse roles could equivalently be described in terms of changes in the salience of external social roles causing changes in the discourse rules that are in play at any one time. We would then simply note that the salience of any social role ends with the end of a conversation. But, this would make talk about role assignment impossible, since we would no longer have roles that could be 'assigned' in the conversation, only saliences that could be altered. The term assignment would then need to be reserved for classical performatives that have effects beyond the conversation. This terminological change would also not change any of the predictions of the theory, only making it more cumbersome to explain. It therefore seems rather easier to adopt the terminology of discourse roles, while recognising that the assignment of a discourse role is equivalent to the external social role becoming salient. Another way to say this is that the term 'discourse role assignment' is syntactic sugar for the phrase 'an increase in salience of an external social role during discourse'. We thank an anonymous referee for making points that forced and enabled us to articulate this equivalence.

[18] Discourse role refers to a short-term social role that only exists for the duration of a particular dialogue. We will use social role as the short-hand for external social role, i.e. any social role that can be ascribed to the person for a period beyond the duration of the dialogue in question. Clearly some social roles will not persist for very long beyond the duration of the conversation, but the distinction stands.

who regard themselves as dominant. Subsequently, because he is acknowledged as a police officer, he is talked to in a different way and is expected to give different types of responses, both in content and tone.

Just as performatives can be used to assign high or low power social roles, so an act of assigning a new discourse role can be used to shape the power dynamics in a conversation. Discourse role assignment is thus a meta-move in the conversational game in that it changes the discourse rules. We saw this in the excerpts above. Mr Tibbs changes his discourse role, by providing the information that his long-term social role is a police officer. At this point, the power balance in the conversation shifts.

Discourse roles are not solitary. A role assigned to one participant typically presupposes complementary roles, to be adopted by other participants. Thus, by entering into a role, a speaker can invite or compel other participants to take on new roles. When Mr. Tibbs enters, the police chief assumes the role of interrogator, thus Mr. Tibbs is compelled to take on the role of the suspect. This first role pairing is complemented by a second, that Mr. Tibbs is a member of a subordinate racial group, and the police chief a member of the dominant racial group. Later, Mr. Tibbs takes on the role of police officer, thus influencing the police chief to take on a more colleague-like role. It is clear from this example that relative power in a dialogue changes when discourse roles change.

The notion of a discourse role is also useful when considering how we talk about non-participants. Social roles come in and out of salience in the same way. Talk about an absent person is still determined in large part by these short-term salient roles. Compare two dialogues:

(8)   *Sue: Johnny is a good and honest boy.*
      Bob: Johnny says he didn't take the biscuits.
(9)   *Sue: Johnny is a thief and a liar.*
      Bob: Johnny says he didn't take the biscuits.

The second utterance in each dialogue is interpreted differently because of the first utterance. Once Johnny has been cast in a role, it influences how utterances about him are interpreted. Either Johnny was a thief and a liar, or he wasn't, both before and after the dialogue. But if the information introduced is new to Bob and other participants, it will change the way utterances about Johnny are interpreted only after the role is introduced. Thus, Johnny's role as discussed in the discourse changes at that point. If Johnny were to enter the room after this point, he would be treated differently depending on whether the start of the conversation was (8) or (9). So at that point, he picks up the status with its discourse rules that were determined for him by a role assignment made in the discourse prior to his entering the room.[19]

---

[19] Later on we will describe how one way of describing the underlying mechanism of role assignment is that of modification of the conversational score. Another way would be to think about this as altering the beliefs in the common ground. So that when we allocate Johnny the role of liar we are altering beliefs about what Johnny is likely to do in certain situations, together with beliefs about how Johnny's utterances ought to be interpreted. The purpose of the roles of "liar" and "thief" is that each puts a set of beliefs on the common ground in one go. It is this idea of roles efficiently indexing many facts, norms and rules that is important. For the purposes of the theory as we describe it here the common ground explanation and the conversational score explanation can be seen as flip sides of the same mechanism for implementing discourse roles and role assignments.

Thus, it seems sensible to think about all such role allocations in discourse as allocating discourse roles, whether or not the person is present. If the person enters the conversation they will be governed by the rules associated with that discourse role, as earlier allocated. However, if the person is not present at any point, the role allocated will still determine how they are spoken about during the conversation, and may affect their social role after the conversation has ended. For example, suppose that Bob didn't previously believe Johnny to be a thief and a liar. If he believes Sue he will certainly treat statements Johnny makes during future conversations with him in a different way than before. This ability for non-participants to have discourse roles, and for those to affect the external social roles of those non-participants post the conversation is important. It will serve as a building block in our later consideration of whether role assignment can account for the effects of both second and third person uses of slurs.

Having established what a discourse role is, we now consider how a slurring utterance alters power relations via speech-acts related to discourse roles. In particular, are slurs oppressive speech?

5 Slurs and oppressive speech

McKinnon (1993) and Langton (1993) proposed that some types of speech, such as pornography, can not only subordinate, but are *constitutive of subordination*. In other words, the illocutionary act performed *is* an act of subordination, rather than merely *causing* subordination through its perlocutionary effects. This position has been attacked on the grounds that the speaker must have authority for the act to fire, and that this is not obviously the case for the kinds of speech they want to so classify. A variety of counter arguments have been marshalled. Langton (1998) proposed that speakers have authority if they already have local control over aspects of the lives of the targets. Maitra (2012), in extending the idea of oppressive speech to hate speech, proposed that the required authority is licensed by the silent assent of the audience. To see a problem with the second approach consider two utterances:

(10) *The neighbourhood has really gone downhill since those spics moved in.*
(11) *Shut up, you fucking faggot!*

While Maitra's argument can explain (10) it cannot solve the authority problem for (11). Suppose there is no audience other than the target for (11). If the speaker is to have authority, then according to Maitra, they must wait on the silence of the target. But, until subordinated, the target can speak out, being on equal terms with the speaker. Thus, the speaker cannot wait for the silence of the target, for surely any rational target would speak out. But targets often don't speak out.[20] Because of this,

---

[20] Nielsen (2012) and West (2012).

Maitra's argument cannot show that one-to-one uses of slurs constitute oppressive speech because she can't solve the authority problem.[21]

McGowan (2009, 2012) argues that, in fact, authority is not needed. To show this she builds on Lewis's (1979) idea of the conversational score. This comprises entities that summarise and are necessary to make sense of the utterances so far in a conversation. Examples of the latter include the relative salience of entities brought under discussion, points of reference, standards of precision in description, relevant possibilities, facts about the world changed by performative utterances, and plans. Lewis (1979: 347) points out that conversational score evolves so as to make the utterances that have occurred "correct play". This is achieved by various "rules of accommodation". In turn, the conversational score and the rules of conversational kinematics together determine what counts as correct play. So an utterance that requires a rule of accommodation changes the conversational score, and thereby changes what is conversationally permissible.

Because it changes what is conversationally permissible, McGowan (2003, 2004) says that any such utterance is a "conversational exercitive". She argues that conversational exercitives don't need to display their communicative intentions, nor have them understood by participants, in order to work, and so they can be what she terms "covert". Critically, she claims that conversational exercitives do not require authority in the way that standard exercitives do. In later work (2009) she extends this idea of a covert exercitive to include any move in any rule (norm) governed activity, i.e. any move in any game changes the state of the game, thus determining the permissible next moves.

Permissible moves are simply defined by the rules of correct play. McGowan splits rules governing games into two classes: g-rules (general rules), and s-rules (specific rules). While g-rules always hold, s-rules can be changed by moves in the game. She then claims that we can think of oppression as a rule governed activity. So, if a verbal move enacts an s-rule within the oppression game then it is an exercitive within the game of oppression, i.e. it changes what is a permissible move in the oppression game. She suggests that sexist and racist speech falls into this category. So sexist and racist speech is constitutive of a move within the game of oppression.

We can use this to explain slurring utterances. We first add discourse roles as being possible entities in the conversational score. These determine what is correct play in the conversation. When a speaker slurs the target they assign them a subordinate discourse role. Because of accommodation the conversational score updates to include this assignment: the target now has the discourse role. This in turn changes rules of interpretation, and rules of permissibility about what the target can say. Thus, when a discourse role is assigned it adds these as s-rules to the conversation. So we first extend Lewis's notion of conversational score to include discourse roles, and then utilise McGowan's notion of a conversational exercitive to give a mechanism by which the discourse roles are changed. In fact the speaker

---

[21] We can think of (11) as a weapon use—an assault that employs its affective effects to subdue the target, thus creating silence, thereby enabling the discourse role assignment to fire by demonstrating power.

imposes not just one, but a complementary pair of discourse roles on the conversational score: the speaker takes on a powerful, dominant role, while assigning the target a subordinate role. Both have associated s-rules.

There is however, a slight difference between the examples of the s-rules that McGowan gives, and the s-rules enacted by a discourse role assignment. This is because the pair of discourse role assignments enacts different s-rules for each participant. In other examples of conversational score, including the oppressive speech examples that McGowan gives, the s-rules alter what is permissible for all participants equally. Roles give a way that s-rules are specific for particular participants. Thus, the s-rules for the dominant role permit the use of types of utterances directed at the subordinate role that would never be permitted in a conversation between participants with equal status. Conversely, the s-rules for the subordinate role restrict the types of allowable utterances for the subordinate role. Because of this the act functions as a mechanism of *unjust and forceful discourse power grab*.

Does this bring light to bear on the debate on the constitutive or causal nature of oppressive speech? It would seem natural, given our division between discourse roles and external social roles, to say that slurs are constitutive of oppressive speech with respect to the discourse roles, because they have the illocutionary force of assigning a subordinate role for the purposes and duration of the discourse. So, in this sense the theory clearly supports the notion that slurs *constitute* oppression in the current social (conversational) interaction.

But what of the effect on the long-term social role external to the discourse? The question is more open here. There are two options. On the one hand, we could maintain that a slurring utterance is constitutive of oppression beyond the current discourse. On the other hand, we could simply say that effects beyond the conversation are causal. To decide between the two we would need to decide what it would mean for the slurring speech-act to be constitutive of oppression beyond the current dialogue. If oppression within the dialogue is the allocation of a subordinate short-term social role (discourse role), then it would follow that oppression beyond the dialogue would be the allocation of a subordinate long-term social role. That, in other words, the target would in all social interactions thereafter, until some other action released them, be in a subordinate social role relative to the speaker. It follows that in order to answer the constitutive versus causal debate in oppressive speech, that we would have to come to an answer as to whether this is plausibly the case or not.

Although we cannot give a full answer in this paper, we sketch the following three positions: *local constitutive (LC)*, *constrained distal constitutive (CDC)* and *universal distal constitutive (UDC)*. Local constitutive is simply the view that the speech-act is oppressive within the current (local) social interaction, i.e. within the current discourse. Constrained distal constitutive is the view that the role assignment persists beyond the current discourse, but that this is constrained to future social interactions with the speaker and aligned audience members. Universal distal constitutive would be the position that all future social interactions between the target and others would be governed by the role assignment. It seems obvious that UDC is false. It cannot be the case that if someone calls a person a slur, that this

triggers a role assignment that persists forever. Choosing between LC and CDC is more complex, since one can easily imagine that some theorists will want to plump for CDC, to maintain that the oppression is constitutive beyond the current social interaction. But equally the LC position has merit, since it follows uncontroversially if the model of this paper is accepted. The arguments are too complex to consider here. But, without choosing between CDC and LC, we would maintain that either satisfies the goal of defining a type of speech (slurring) as constituting oppression. In the latter case (LC) the consequence would be that oppression beyond the current social interaction (discourse) is *caused* by the slurring utterance.

Whether slurs constitute (CDC) or cause (LC) oppression beyond the dialogue, we need to make it clearer how it happens. We now turn to this. We will argue that it is clear that at least some of the effects are causal, i.e. perlocutionary.

6 How slurs contribute to oppression

We've argued that slurring acts are moves in a discourse that seek to shift power from the target to the speaker. If a slurring act only affected the current discourse, its effects would end when the discourse ends. But they don't. So, how do they affect the social roles of targets in the real world?

Just as the discourse role is influenced by the social roles of the participants, so the social roles of the participants are influenced by the discourse. The discourse role feeds back into the social role. This feedback is achieved via perlocutionary effects of the slurring utterances.

First, these include changes to the attitudes of audience members and to their dispositions about how to treat members of the target group. There is evidence that social circumstances influence the extent to which people discriminate, and can override an individual's declared ideology. Warner and DeFleur (1969) shows that those of low racial prejudice are more likely to act in a discriminatory manner if they are members of a society in which there is a significant number of bigots, and they know that their actions will be broadcast. Why? They factor in the social cost of disapproval from a dominantly bigoted community. In predominantly non-bigoted societies the reverse is true: bigots hide their beliefs so as to avoid social disapproval. Thus, when a bigot uses slur words, they make it more permissible for others within their social circle to give voice to their bigotry. If there are sufficient bigots they create a pressure for even non-bigots to discriminate. In this way, the perlocutionary effect of a slurring use is to increase oppression by incrementally increasing the social pressure for, and acceptability of, oppressive acts.

But these effects may have different results in different circumstances. So that racial slurring, for example, in a racist society versus a non-racist society will have different effects in inducing, or failing to induce, others to discriminatory acts. This difference, in effect based on circumstance, can be explained in terms of Kukla (2014) and Kukla and Lance (2009), Lance and Kukla (2013) who describe how speech acts shift norms. They introduce the notion of uptake. In this, the pragmatic context within which the utterance is made determines whether uptake is achieved. They refer to the "output" of a speech act as the set of normative statuses it

institutes. They emphasise that "a speech act can be, in virtue of its uptake, a different speech act than it would typically be, given its social context and standard discursive conventions" (Kukla 2014: 3). To provide a complete account of how slurs shift norms and how their effects change in different social circumstances we will likely need to draw on such notions.

Second, by slurring the bigot shows others the power they can acquire. On the power theory outlined here, the bigot is not talking about power, they are demonstrating power. The speaker acquires discourse power. We posit that this is emotionally appealing to audience members—who are not members of the target group—who feel less powerful than they would like. They see that they can accrue power to themselves by using a slur. Thus, a perlocutionary effect is to make others desire the power the bigot has grabbed. Increasing desire is different to increasing acceptability. Both are required for audience members to join the side of the bigot.

A slurring act also has perlocutionary effects on the target. First, targets are affected as after a threat of physical assault. As Lawrence (1993) writes: "One of my students, a white, gay male … told a story of being called "faggot" by a man on a subway. His description included all of the speech inhibiting elements we have noted previously. He found himself in a state of semi-shock, nauseous, dizzy, unable to muster the witty, sarcastic, articulate rejoinder he was accustomed to making. He was instantly aware of the recent spate of gay bashing in San Francisco, and that many of incidents had escalated from verbal encounters."

This demonstrates that the effects on the target are those of the fight-flight response, and that this is because the slurring use carries an implicit threat of violence. This in turn, Lawrence noted, led to a speech inhibiting effect, to silencing. Targets are more likely to be the victims of violence, either at the time of slurring, or afterwards. Imagine a racist policeman who uses a racial slur while arresting a member of an ethnic minority. Slurs associated with a history of violence implicate the potential for this violence to be re-enacted. If you were an African-American, and a policeman stopped you in your car and called you "Nigger", would you merely feel offended, or would you also fear for your physical safety?

This ability to incite to action is shown by Tirrell (2012: 174), who argues that deeply derogatory words are "action-engendering" and "can be part and parcel of genocide, not only an antecedent of it". When Tutsi were labeled "inyenzi' (cockroach) or 'inzoka' (snake), this granted permission to hate and a mandate to kill. As she says, "understanding these speech-acts helps to illuminate the important ways that power is enacted through discourse, how speech-acts can prepare the way for physical and material acts, and how speech generates permissions for actions hitherto uncountenanced."

This is also recognised within the US law enforcement community. For example: "San Francisco's public defender, Jeff Adachi, sharply criticized three city police officers for text messages they exchanged … using racial slurs … 'That's the kind of mentality that tells you it's OK to shoot, OK to kill, OK to arrest' people of color, Adachi told reporters" (Chang and Tchekmedyian 2016). This is evidenced by studies that show an association between increased rates of internet searches using racial slurs for African-Americans and various social effects, including an increased likelihood of voting against a black politician (Stephens-Davidowitz 2014), an

increased likelihood of African-Americans being shot by the police while unarmed relative to white people (Ross 2015), and increased mortality rates among African-Americans (Chae et al. 2015).

Thus, perlocutionary effects causing oppression beyond the discourse include: emotional injury, implicit threat of physical injury, silencing, increased permissibility and/or pressure for other oppressive acts, and increased desire to act oppressively so as to gain power.

7 Explaining variable offence

The theory set out so far has several elements. It is now worth running through these, to see quickly how they lock together. First, we have the idea of both short term-social roles (discourse roles) and of longer-term social roles. Second, we have the idea that these roles often come in pairs, i.e. that they make most sense when defined in relation to one another. Third, we noted that some pairs of long-term social roles are associated with a power imbalance. Fourth, that some such power imbalances are unjust. Fifth, we then have the idea that in a social interaction, such as a conversation, a social role can be assigned. This can be a long-term social role, such as achieved by a regular exercitive—"I declare you man and wife"—or it can be a role that exists for the purposes of the discourse. Sixth, we then considered a candidate mechanism for the assignment of such roles, and found that the notion of a conversational game from Lewis, augmented by the notion of a conversational exercitive from McGowan, seems to fit the bill. Finally, we laid out the range of positions that seem plausible with regard to whether discourse role assignment of a subordinate role could be seen as either constitutive or causal of oppression We refer to this set of seven linked hypotheses as the 'role and unjust power' theory of oppressive speech, or *role-power theory* for short.

This leads us to state the prediction that the theory makes about variable offence. This is that *the degree of offence taken by an audience member in relation to a slurring speech-act is in proportion to the perceived unjustness—according to that audience member—of the power imbalance that it plausibly achieves*. Let's unpack this.

First, we need to understand how the power imbalance that it [the slurring speech-act] plausibly achieves can vary across terms. The power imbalance created is between the short-term social roles, i.e. the discourse roles allocated to each participant. Earlier we saw that discourse roles draw upon longer-term social roles that exist externally to the current discourse. An oppressive slur term can be used to allocate a pair of discourse roles, where these roles draw upon a pair of roles that had a history of oppression. Thus, the greater the degree of oppression associated with the long-term historical roles, the greater the offence likely to be caused. This would neatly explain across group word variation, e.g. why a racial slur directed at African-Americans are generally considered to be more offensive than those directed at other ethnic groups. So that "Nigger" is typically considered by a non-racist to be more offensive than "Chink".

The limiting case would thus be slur terms associated with a privileged group, such as "Limey", "Yankee" or "Honkey". These are generally accepted as non-offensive. In this case it could be said that these groups have, in relation to other groups, actually been perpetuators of oppression. In such a case what would we expect the theory to predict? It would be reasonable to suppose that if a word evokes no long-term social role for the target subject to unjust oppression, then there can be no offence. This is what we typically find. It is, of course, possible for the utterance to be used with a contemptuous tone, and thus the speaker's expression of contempt via that tone to cause offence, but there is no additional offence caused by the use of the particular word.

What about offence variation across words used to refer to the same group? We know that "nigger" is widely considered more offensive than "spook". Does this cause a problem? If both words reference the same role, then surely they reference the same history of oppression. There are two positions that could be taken to explain this. The first is based on the way that the historical use of each term varies, so that "nigger" first had a five-fold upswing in its printed use between 1848 and 1862 when African-Americans were suffering the extreme oppression of slavery.[22] "Spook", on the other hand, does not appear as a racial slur until the twentieth century, and could thus not be directly associated with the roles involved in slavery. So, on this view, "nigger" and "spook" would actually index different roles, because they index different (if overlapping) periods in the history of oppression.

A different explanation would be that words tend to be associated with the history of oppression based on the frequency of their use. This account would also be consistent with data on use frequency, as "spook" (not even removing the non-racial uses of the term) occurred, even during its history as a racial term, no more than 1/10th as often as "nigger".[23] This explanation would predict that offence would rise with more frequent use. An extension of this second explanation would be that since slur terms such as "spook" are now archaic, memory of them gradually falls away altogether, so that they cause no offence to those who do not know their earlier use. Thus zero use eventually leads to zero offence.

So, VT1 in general is a function of two factors. First, it depends on the extent to which the slur word evokes (via association or familiarity of slurring use) a pair of roles from the history of oppression. Second, it depends on the extent of the unjust power imbalance present in the external social roles so indexed.

What about use-variation (VT2)? This covers the variation in offence caused according to who the speaker is. So, for example, it is quite common now—although it was not always—for African-American comedians to use the term "nigger" without offence, but extremely rare for a white comedian to be able to do so. Can we provide an explanation of this? The answer lies in the dual role

---

[22] Source: Google Books n-gram analysis. All the statements regarding word frequency and periods of use are taken from this source.

[23] Google n-gram shows that the usage of "spook" is between ten and sixteen times less frequent than "nigger" in the corpus of American English (2012). While this doesn't track non-written uses, it is likely that verbal and non-verbal uses track each other. It should also be noted that the frequency of "spook" as a racial slur is even less frequent, because there are two other non-slur meanings of the term.

assignment that occurs. If the speaker cannot plausibly fit the role, then role assignment is much less likely to work. It is a felicity condition for the slurring act. If the speaker does not fit the role of the oppressor, then offence can be avoided. Discourse roles often inherit from external social roles. When a discourse role is assigned, it presupposes (typically via the associated external social roles) certain facts about the speaker and target. While the conversational score updates using accommodation, i.e. to make a move correct play, there are limits: accommodation isn't infinitely elastic. Thus, if the presupposed facts about speaker or target do not fit the actual facts, then accommodation breaks. Thus, if the speaker is gay and the target is heterosexual, the use of the slur "fag" presupposes facts (that the target is gay and the speaker is prejudiced against gay people) which do not fit their actual social identities.

It may also be that the existing discourse roles of the participants contain rules that prevent one of the speakers from subordinating the other via slurring. This is typically the case when the speaker is subordinate to the target at the start of the discourse. Thus, take Lewis's example of the master and slave. The normative social interaction between these two is governed by the social roles of master and slave, so the conversation starts with the roles of master and slave as the default discourse roles. If the slave attempts to subordinate the master by slurring, this fails because the discourse rules already governing the conversation prevent the slave from subordinating the master. Thus, again, accommodation has limits. This function of discourse roles limiting the moves that may be made by some speakers, thus maintaining them in a subordinate role, is related to accounts of discursive injustice (Kukla 2014).

But we need to be careful here. Some cases where social identities do not perfectly fit still allow offensive slurring uses. It is the case that in-group uses can sometimes offend, and also that out-group uses can sometimes fail to offend. What causes this inconsistency? We would posit that it is not just group-membership, but a whole variety of other factors that can modulate whether the felicity condition is met. So if there is a plausible power differential between the speaker and target, that can drive offence. Recall the account of an African-American woman being offended when her African-American boss called her a "Nigger" (Gabbat 2013). Other flags, such as tone, familiarity between the participants, appropriateness of context, will all modulate whether the felicity condition is met. Thus, in-group versus out-group is just one part of the felicity conditions. We would suggest that one way to think of these felicity conditions is as additive: if a sufficient number are met, then it is highly likely that the target will be offended. But the number of the sub-conditions that have to be met for offence to be triggered varies across audience members.

What of audience-variation (VT3) more generally? This is where one audience member can be offended by a particular use and another not. In this case the theory says that the offence rises with *the perceived unjustness—according to that audience member—of the power imbalance.* Thus, the contention is that the degree of offence taken by an individual is due to the difference between the power imbalance they *see as appropriate* and the power imbalance they *perceive as being assigned* by the slurring act. The more bigoted the listener is, the less offended they

are, because the more appropriate they see the power imbalance as being. Bigots aren't offended, because they believe it is just that the target group be subordinate.

Power imbalance through role assignment can be shown to thus neatly correlate with all three types of variable offence. But all the uses so far are direct or second-person uses. What happens when a third-person use occurs?[24] To answer this imagine that at a party two people are talking.

(12) Bill: "The neighbourhood has really gone downhill since those spics moved in."
      Sue: "Actually my husband is Hispanic."

Here Sue is offended because Bill is making a statement about Hispanics that she clearly disagrees with. But, how can the role assignment occur if the target isn't present? The clue to this is in the example of Johnny from earlier in the paper. We saw that if the role assignment were accepted, Johnny would be treated according to the role previously assigned to him, on entering the conversation. This is exactly what is happening here. To see that the role assignment really occurs in third-person uses imagine this variation of the above dialogue:

(13) Bill: "The neighbourhood has really gone downhill since those spics moved in."
      Sue: "I know."
      <later on Leila, who is one of their Hispanic neighbours, approaches and joins in the conversation having failed to overhear the earlier remark>

In this case it will be clear that there is now agreement, by Sue, that Bill and Sue will treat Leila according to the role he has allocated of "spic". So although Leila was not present when the utterance was made, she will still be treated differently. Now, of course, it could be the case that Bill and Sue in this second dialogue are closet racists, that they don't want to explicitly let Leila know what they think of her, but even in this case they will treat her differently, and each of Bill and Sue will understand the behaviour of the other towards Leila in terms of their view of how someone with the role of "spic" is to be treated. By knowing that they are in agreement about Leila's role they will be able to oppress in concert with one another. Thus, the third-person use does affect the social role of the target in the interaction even though they were not present when the utterance was made. Clearly, however, Leila was not offended, because she was not aware of the role assignment. But Bill's role assignment is picked-up by Leila when she enters the conversation. We can see this because we can see that Sue would still know, when Leila arrived, that Bill would treat Leila as a "spic".

We will now finish by sketching how power through role assignment can also account for silencing and appropriation.

---

[24] Slurs among out-group members are in fact quite common. There are many possible reasons for this. As Nunberg (forthcoming) points out, intra-bigot talk serves to "create solidarity in a common sense of resentment or superiority; to enjoy the complicit schoolyard-variety naughtiness in using forbidden words, particularly in the form of racial or homophobic humour; or to underscore the normative values of the group".

8 Explaining silencing and appropriation

In this final section we will very briefly sketch how two other phenomena associated with slurs might be explained. These are silencing and re-appropriation or reclaim. A full treatment requires a different paper, but a sketch can be instructive.

Silencing is actually a pair of phenomena. First, the target feels intimidated to respond. Second, even if they do respond their responses cannot completely undo the effects of the slur, and can be ignored. Let us take the example of the student of Lawrence (1993), who suffered the perlocutionary effects of being in a state of semi-shock, and unable to muster a rejoinder. This suggests not that the victim accepts their role as subordinate, but that the role pair includes the potential for violence by the speaker. This violence is part of what is seen as appropriate treatment according to the bigot. It doesn't matter whether the target disagrees. Perpetrating a violent attack is part of the role of the oppressor, and thus via the dual role assignment they signal they may carry this out. So this is both constitutive and causal. The utterance constitutes a threat of violence by reference to the roles assigned, and one of the perlocutionary effects is the fear of the target that self-silences to try to self-preserve.

Furthermore, in constitutive terms the discourse role also includes rules that permit participants to ignore future utterances by the target, including those that disable objections from the target to the assignment itself. So, take the example of a speaker calling a woman a "slut". An objection she makes could be disabled by saying that it is part of her role to deny this, and that thus her denial can be safely ignored. Or imagine an African-American man arrested by a racist policeman who targets him with a slur: if he objects, his objections can be used as evidence of his criminality—resisting arrest, talking back to a police officer. In this way, his perfectly justified response is turned against him. So in the first case the target does not speak, and in the second case they will not be heard. Both are demonstrations of power.

What about appropriation? This is a complex phenomenon. We saw that, if the speaker mis-matches the role of the oppressor,[25] the act of discourse role assignment doesn't necessarily fire, because the felicity conditions aren't met. Thus, when a member of the target group uses a slur word, the felicity conditions for the role assignment don't automatically fire because their group-membership doesn't match that of the historical oppressor group. Because of this failure, the slur word is freed, enabling appropriation as a possibility. This doesn't mean that all in-group uses enable appropriation. Cases of anger, contempt, or self-loathing may still trigger the felicity conditions. Instead, the space for appropriation must be created by repeated uses where the felicity conditions are violated.

Appropriation enables members of the target group to put the slur term to better use.[26] So when an African-American greets another African-American with

---

[25] Matches can be accommodating, e.g. an ethnically Asian policeman in San Francisco used slurs against African-Americans.

[26] Bianchi (2014) and Herbert (2015), among others.

"Nigger", a different pair of roles can be assigned—roles with equal discourse rights and camaraderie. This is not universal; it is perfectly possible for in-group uses to offend if some of the felicity conditions are met. Additionally, it is not the case that anyone can use an appropriated term. A friendly white speaker of "Nigger" will likely offend, insofar as their group-membership fulfils the felicity conditions for a speech-act that assigns roles with an unjust power imbalance. So the two uses (in-group and out-group) have the potential to assign different pairs of roles. Thus, all we have established is that by breaking the felicity conditions in some cases, in-group uses can create a space within which re-appropriation can begin. This does not say how that re-appropriation occurs, and this should be a topic for further consideration.

Finally, the idea of unjust power is what draws the distinction between oppressive slurs and other derogatory expressions, such as pejoratives. Take "dickhead": it expresses contempt, but it doesn't oppress by allocating a low power role that draws on a history of oppression. It is certainly the case that the target of a pejorative such as "dickhead" may be offended. It is also true, however, that "dickhead" is a term with some kind of role. We think of a person who does socially inappropriate things and doesn't care about their consequences. But this role is not an oppressed role in society. This distinction between pejoratives and oppressive slurs thus rests on the notion of power, rather than the notion of role. It can of course be said that in calling someone a "dickhead" the speaker is performing a conversational exercitive, and all such exercitives can leave the target with a sense that something has been done to them in the conversation that they did not want or permit, and this is indeed an example of force being applied in dialogue. In this sense, again, a pejorative and a slur are similar in that they take the initiative. But there is a difference between this and the unjustness of the power associated with the roles assigned. A pejorative such as "dickhead" does not have a natural paired role that oppresses them.

9 Conclusion

Unjust power is an important dimension along which offence varies for slurring utterances. I've argued that this explains three kinds of variability other accounts struggle with. It also explains silencing and the possibility of appropriation. There are other dimensions required to explain.

The shape of the solution is as follows. Power relations, and thus oppression, are captured by social roles. Discourse roles inherit from social roles, and reflect the power of those social roles in their discourse rules. Discourse roles are part of the conversational score. They thus determine the rules of conversational kinematics via their associated discourse rules. Discourse roles can be assigned by a variety of utterances, which work as conversational versions of Austinian exercitives. A slurring utterance assigns a pair of discourse roles, between which there is an unjust power imbalance. This affects power in the dialogue. The slurring move also has perlocutionary effects on both target and audience that cause the spread and intensification of oppression.

This theory relates what people do with language to the social structure within which language use is situated. The primary theoretical innovation has been to extend Lewis's notion of a scorekeeping in a language game with the notion of discourse roles which can be changed by dialogue moves. The idea of related discourse and social roles that feed off one another is, however, not restricted to slurring, and may find application in other problems. For example, when non-slurring speech assigns roles, such as the case of Johnny being called a liar, it has similar effects, as we saw. Another, more topical example is the use by Donald Trump of names that assign roles to his political opponents: "Crooked Hillary", "Little Marco" and "Lyin' Ted". These are not slurs, but deliberately turn names into role allocations.

References

Anderson, E. (2012). *The imperative of integration*. Princeton: Princeton 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 words. *Analytic Philosophy, 54*(3), 350–363.

Bach, K. (Forthcoming). Mean and nasty talk: On the semantics and pragmatics of slurs. In D. Sosa (Ed.), *Bad words*. Oxford: Oxford University Press.

Barker, S. J. (2000). Is value content a component of conventional implicature? *Analysis, 60*(267), 268–279.

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

Bolinger, R. J. (2015). The pragmatics of slurs. *Noûs, 49*(1), 1–24.

Burge, T. (1986). Individualism and psychology. *Philosophical Review, 95,* 3–45.

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

Chae, D. H., Clouston, S., Hatzenbuehler, M. L., Kramer, M. R., Cooper, H. L. F., Wilson, S. M., et al. (2015). Association between an internet-based measure of area racism and black mortality. *PloS One*. doi:10.1371/journal.pone.0122963.

Chang, C., & Tchekmedyian, A. (2016). Racial slurs by law enforcement are a legacy that's becoming more unacceptable. *Los Angeles Times, May 3 2016*.

Copp, D. (2001). Realist expressivism: A neglected option for moral realism. *Social Philosophy and Policy, 18,* 1–43.

Cudd, A. (2006). *Analyzing oppression*. Oxford: Oxford University Press.

Dahl, R. (1957). The concept of power. *Behavioral Science, 2,* 201–215.

Gabbat, A. (2013). Black woman wins civil case against black manager who used the N-word. *The Guardian*, Wednesday 4 September. www.theguardian.com/world/2013/sep/04/black-woman-n-word-court-case.

Goffman, E. (1961). *Encounters: Two studies in the sociology of interaction*. Indianapolis: Bobbs-Merrill.

Haslanger, S. (2012). *Resisting reality: Social construction and social critique*. Oxford: Oxford University Press.

Herbert, C. (2015). Precarious projects: The performative structure of appropriation. *Language Sciences, 52,* 131–138.

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

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

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

Hornsby, J. (1995). Disempowered speech. *Philosophical Topics, 23*(2), 127–147.

Hornsby, J. (2001). Meaning and uselessness. *Midwest Studies in Philosophy, 25*(1), 128–141.

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

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

Jeshion, R. (Forthcoming). Slurs, dehumanization, and the expression of contempt. In D. Sosa (Ed.), *Bad words*. Oxford: Oxford University Press.

Kukla, R. (2014). Performative force, convention, and discursive injustice. *Hypatia, 29*(2), 440–457.

Kukla, R., & Lance, M. (2009). *'Yo!' and 'Lo!': The pragmatic topography of the space of reasons*. Cambridge, MA: Harvard University Press.

Lance, M., & Kukla, R. (2013). Leave the gun; take the cannoli! The pragmatic topography of second-person calls. *Ethics, 123,* 456–478.

Langton, R. (1993). Speech acts and unspeakable acts. *Philosophy and Public Affairs, 22*(4), 293–330.

Langton, R. (1998). Subordination, silence, and pornography's authority. In R. Post (Ed.), *Censorship and silencing: Practices of cultural regulation* (pp. 261–284). Oxford: Oxford University Press.

Langton, R. (2012). Beyond belief: Pragmatics in hate speech and pornography. In I. Maitra & M. K. McGowan (Eds.), *Speech and harm: Controversies over free speech* (pp. 94–120). Ocford: Oxford University Press.

Lawrence, C. R. (1993). If he hollers let him go. In M. Matsuda, C. R. Lawrence, R. Delgado, & K. Williams Crenshaw (Eds.), *Words that wound*. Boulder, CO: Westview Press.

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

Lukes, S. (1974). *Power: A radical view*. London: Macmillan.

Lycan, W. (2015). Slurs and lexical presumption. *Language Sciences, 52,* 3–11.

Maitra, I. (2009). Silencing speech. *Canadian Journal of Philosophy, 39*(2), 309–338.

Maitra, I. (2012). Subordinating speech. In I. Maitra & M. K. McGowan (Eds.), *Speech and harm: Controversies over free speech* (pp. 94–120). Oxford: Oxford University Press.

Mallon, R. (2003). Social construction, social roles, and stability. In F. Schmitt (Ed.), *Socializing metaphysics: The nature of social reality* (pp. 327–354). Lanham, MD: Rowman and Littlefield.

Matsuda, M., Lawrence, C. R., Delgado, R., & Williams Crenshaw, K. (1993). *Words that wound*. Boulder, CO: Westview Press.

McGowan, M. K. (2003). Conversational exercitives and the force of pornography. *Philosophy and Public Affairs, 31,* 155–189.

McGowan, M. K. (2004). Conversational exercitives: Something else we do with our words. *Linguistics and Philosophy, 27,* 93–111.

McGowan, M. K. (2009). Oppressive speech. *Australasian Journal of Philosophy, 87*(3), 389–407.

McGowan, M. K. (2012). On "whites only" signs and racist hate speech: Verbal acts of racial discrimination. In I. Maitra & M. K. McGowan (Eds.), *Speech and harm: Controversies over free speech* (pp. 222–250). Oxford: Oxford University Press.

McKinnon, C. (1993). *Only words*. Cambridge, MA: Harvard University Press.

Nielsen, L. B. (2012). Power in public: Reactions, responses, and resistance to offensive public speech. In I. Maitra & M. K. McGowan (Eds.), *Speech and harm: Controversies over free speech* (pp. 148–173). Oxford: Oxford University Press.

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

Popa-Wyatt, M. (2016). Not all slurs are equal. *Phenomenology and Mind, 11,* 150–156.

Potts, C. (2007). The expressive dimension. *Theoretical Linguistics, 33*(2), 165–197.

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

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

Ross, C. T. (2015). A multi-level Bayesian analysis of racial bias in police shootings at the county-level in the United States, 2011–2014. *PloS One*. doi:10.1371/journal.pone.0141854.

Saka, P. (2007). *How to think about meaning*. Berlin: Springer.

Schlenker, P. (2007). Expressive presuppositions. *Linguistics, 33*(2), 237–245.

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

Stephens-Davidowitz, S. (2014). The cost of racial animus on a black candidate: Evidence using Google search data. *Journal of Public Economics, 118,* 26–40.

Swanson, E. (ms). Slurs and ideologies. In *Ideology*. Oxford: Oxford University Press.

Tirrell, L. (1999). Derogatory words: Racisim, sexism, and the inferential theory of meaning. In C. Hendricks & K. Oliver (Eds.), *Language and liberation: Feminism, philosophy, and language* (pp. 41–79). Albany, NY: SUNY Press.

Tirrell, L. (2012). Genocidal language games. In I. Maitra & M. K. McGowan (Eds.), *Speech and harm: Controversies over Free Speech* (pp. 174–221). Oxford: Oxford University Press.

Warner, L., & DeFleur, M. L. (1969). Attitude as an interactional concept: Social constraint and social distance as intervening variables between attitudes and action. *American Sociological Review, 34,* 153–169.

Weber, M. (1978). *Economy and society: An outline of interpretive sociology* (trans: Ephraim Fischoff et al.). Berkeley, CA: University of California Press.

West, C. (2012). Words that silence? Freedom of expression and racist hate speech. In I. Maitra & M. K. McGowan (Eds.), *Speech and harm: Controversies over free speech* (pp. 222–250). Oxford: Oxford University Press.

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

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.

Young, M. (1990). Five faces of oppression. *Justice and the politics of difference* (pp. 39–65). Princeton: Princeton University Press.