

Note: this is an article whose final and definitive version will be published in the *Australasian Journal of Philosophy* in 2013; the *Australasian Journal of Philosophy* is available online at: <http://www.tandf.co.uk/journals/>.

Un-ringing the bell: McGowan on oppressive speech and the asymmetric pliability of conversations

Robert Mark Simpson

Abstract: In recent work Mary Kate McGowan presents an account of oppressive speech inspired by David Lewis's analysis of conversational kinematics. Speech can effect identity-based oppression, McGowan argues, by altering 'the conversational score' – which is to say, roughly, that it can introduce presuppositions and expectations into a conversation, and thus determine what sort of subsequent conversational 'moves' are apt, correct, felicitous, etc. – in a manner that oppresses members of a certain group (e.g. because the suppositions and expectations derogate or demean members of that group). In keeping with the Lewisian picture, McGowan stresses the *asymmetric pliability* of conversational scores. She argues that it is easier to introduce (for example) sexist presuppositions and expectations into a conversation than it is to remove them. Responding to a sexist remark, she thus suggests, is like trying to 'unring a bell'. I begin by situating McGowan's work in the wider literature on speech and social hierarchy, and explaining how her account of oppressive speech improves upon the work of others in its explication of the relationship between individuals' verbal conduct and structurally oppressive social arrangements. I then propose an explanation and supportive elaboration of McGowan's claims about the asymmetric pliability of conversations involving identity-oppressive speech. Rather than regarding such asymmetry as a *sui generis* phenomenon, I show how we can understand it as a consequence of a more general asymmetry between making things salient and *un-salient* in speech, and I show how this asymmetry also operates in various cases that interested Lewis.

1. Introduction

Recently there has been a concerted effort among certain philosophers, legal theorists, and political theorists, to show that speech and other expressive conduct should be central to our attempts to understand and explain identity-based oppression and social hierarchies. Mary Kate McGowan has made notable contributions in this area, by employing David Lewis's analysis of conversational kinematics to develop an account of how verbal conduct is implicated in identity-based structural oppression (an account of the sort that would, in principle, vindicate McGowan's and others' speech-oriented approach to issues of social equality). McGowan's discussion focuses on the way that casual sexist remarks can 'enact oppressive permissibility facts', not because the speaker has special power or authority as such, but simply as a result of some general features of conversational interaction, features to which Lewis drew our attention. McGowan then argues that there is a crucial asymmetry in the way that identity-oppressive language works, such that it is relatively easy for speakers to enact identity-oppressive permissibility facts in conversation, but relatively difficult for anyone to rescind those permissibility facts.

McGowan's work shares some admirable qualities with other recent, influential writing in this area, most notably in its application of conceptual tools from analytic philosophy of language to a complex set of politically-charged questions about social hierarchy. But in certain respects her work is, so I will argue, an advance on what has come before. My first purpose here is to illuminate the significant strengths of McGowan's analytic apparatus, in particular the way in which her account articulates – in both senses of the word – the nature of the relationship between structurally oppressive social arrangements and verbal conduct that positively contributes to those arrangements. My second aim is to present a sympathetic elaboration of one of McGowan's key claims, namely, her claim that conversations involving identity-oppressive utterances are asymmetrically pliable. I show how we can use the idea of associative schemas to characterise and explain the phenomena in question, such that the alleged asymmetries are understood as a consequence of a more general asymmetry between (verbally) making something salient and (verbally) making something *un*-salient. Throughout I aim to show how McGowan's ideas speak to a wider set of questions, about whether and why we should pay special attention to linguistic considerations in our attempts to understand and combat identity-based social inequality. In both its approach and its substance, McGowan's analysis supports the notion that it is entirely appropriate to focus on speech in relation to these issues.

2. Lewis on conversational score

One of David Lewis's contributions to philosophy of language was the notion of 'conversational score'. For a conversation *C* at time *t*, the conversational score is, roughly, an informal register of the participants' shared presuppositions and shared or respective expectations, which together determine – in accordance with complex but consistently-behaved rules – which conversational moves would constitute 'correct play' in *C* at *t* [Lewis 1979: 344-45]. Lewis devised this notion to aid in his analysis of conversational kinematics. When we compare the rules governing the operation of conversations with the rules of another complex, socially-coordinated activity, such as a baseball game, we see that conversations have a special tendency 'to evolve in such a way as is required in order to make whatever occurs [in a given conversation] count as correct play' [Lewis 1979: 347]. In baseball, on one hand, it is incorrect for a batter to walk after the third ball; in a conversation about the Scottish independence movement, on the other, it is incorrect for person *A* to start talking, apropos of nothing, about how much she loves soft cheeses. But in the conversation, unlike the baseball match, an 'incorrect move' like this one is easily accommodated, so that the activity can carry on without interruption (or without us having occasion to worry about whether we are still, strictly speaking, carrying on the same activity, as we may do in a non-standardly conducted baseball match). It's true that *A*'s conversational gambit can be challenged, e.g. by one of her audience saying: 'What's this about cheese? We were talking politics'. But unless this occurs, the conversational score will update so that comments about cheese become appropriate fodder for discussion, and so that *A*'s comment may be retrospectively counted as correct play [M. S. Green 2000: 466]. In this way conversations are distinctively pliable. If someone wants to change the trajectory of a conversation so that certain ideas are 'in play', or so that certain conversational moves become 'correct play', this purpose can be realised simply by making a suitable interjection.

Lewis was eager to stress, however, that although conversational scores are pliable, they are not *uniformly* pliable. He illustrated this point with examples of conversations whose constituent statements include vague predicates. In a conversation about the relationship between belief and knowledge, person *A* may invoke the concept of certainty: she may say that she is *certain* she owns fewer than fifty pairs of socks. Her interlocutor *B* may then, in a routine sceptical manoeuvre, ask whether *A* is *really* certain of that. And thus, by posing a question that is predicated upon a stringent interpretation of the vague predicate 'certain', *B* may alter the conversational score so that some typical utterances incorporating that term (e.g. 'I'm certain I haven't bought any socks in the last year') will thereafter be incorrect – or at least, conspicuously infelicitous – contributions to the conversation. So far this just looks like another example

of conversational pliability. Lewis's point, however, is that it is more difficult to shift the conversational score in the other direction, so that relatively lax usages of 'certain' can be felicitously invoked in subsequent conversational moves. For some reason (Lewis does not venture an explanation) 'raising of standards goes more smoothly than lowering', and this is true despite the fact that lax usages of vague predicates are generally no less useful or informative than stringent uses [Lewis 1979: 352]. In these sorts of cases conversational scores are not merely pliable, they are *asymmetrically* pliable.¹

3. A language-oriented approach to understanding identity-based oppression

Lewis took himself to be offering general observations about how conversations operate, and if he was right about the generality of his observations we should not be surprised to see the phenomena in question occurring in areas to which Lewis paid little attention. Nevertheless, it may be at least a little surprising to see a Lewisian analysis of conversational score brought to bear in a discussion of identity-based oppression. One might expect that in this arena we would be drawing upon the normative conceptual resources of political, legal, and economic theory, instead of a rather dry – and on its face not politically-oriented – piece of linguistic pragmatics. Those familiar with the relevant literature, however, will know that linguistic preoccupations in the discussion of identity-oppression are not unusual among today's authors, if they ever were. Since the 1980s and the emergence of radical feminist critiques of pornography [e.g. Dworkin 1981; MacKinnon 1984; Dworkin and MacKinnon 1988], and critical race theory [e.g. Delgado 1993; Lawrence 1993; Matsuda 1993], something like the following thesis has often been endorsed: that in order to diagnose the causes and character of identity-based oppression, we must examine the nature and social operations of the speech acts (and other forms of symbolic expression) that are used to condone or promote such oppression, and/or to articulate its underlying ideas and attitudes. *Why* are language and speech taken to be important in understanding identity-oppression? Because of a conviction that, as one author puts it, 'words create the hierarchies and people fill them' [Gilreath 2009: 604]. Stated in more cautious terms, the suggestion is that identity-oppressive speech plays a crucial role

¹ Lewis observes the same structural phenomenon in conversations involving relative modal terms. A says: 'I can't attend the managers' meeting on Wednesday'; B replies: 'You *can't* or you *won't*?', and thus she alters the conversational score so that correct play thereafter requires the use of 'can't' to accord with a more stringent standard (e.g. 'cannot within the realm of what is physically possible') rather than A's initial, less stringent usage. But whereas this change in the conversational score happens quite effortlessly, it is difficult to reverse the score so that a less stringent usage of 'can't' once again becomes correct play.

in the complex, multi-faceted causal web through which identity-oppression *per se* is effected, as opposed to it merely being a peripheral element in that causal nexus, or just an eventual outworking of its operations. According to this line of thought, if we want to understand identity-based oppression – how it works and how it can be combated – we must examine the workings of oppressive speech.

Exactly what a good account of this subject would involve is of course open to debate. Probably the most well-known and well-received recent body of distinctively philosophical work in this area is that produced Rae Langton and Jennifer Hornsby and others following in their wake. These authors adapt the conceptual resources of Austin's [1962] and Searle's [1969] speech act theory, principally, the distinction between (i) the *locutionary content* of an utterance, and (ii) the *illocutionary use* of that utterance (i.e. the act that is performed *in* the utterance of a certain locution), in order to interpret and defend the claims of feminist authors like Catharine MacKinnon and Andrea Dworkin, in particular, their claims that pornography subordinates and silences women. In the hands of Langton and Hornsby et al., the first claim is interpreted as saying that subordination is a kind of partially exercitive and partially verdictive illocutionary act, which involves ranking women as inferior to men, and that certain types of graphic, sexually explicit expression (i.e. pornography) can be used to perform the illocutionary acts in question [see in particular Langton 1993; Langton 1998]. The claim that pornography silences women, then, is interpreted as saying that pornography jeopardises the felicity conditions for *women's* performance of certain important illocutionary acts (e.g. the use of speech to refuse a sexual advance) [see for instance Hornsby 1994; Hornsby and Langton 1998; Maitra 2009].

This interconnected body of literature exhibits a great deal of philosophical acuity, and provides much illuminating analysis of language, and of identity-oppression, and of the relationships that obtain between the two. Whether this kind of work ultimately vindicates the conviction from which it springs – whether it shows that language *is indeed* at the causal or explanatory core of identity-oppression – can only be adjudicated after close critical attention. For some readers, it will be a cause for initial suspicion that the anti-pornography feminists and critical race theorists (and others who adopt a language-oriented approach here) seem to be amenable to a controversial kind of idealist thought in philosophy of language, one which has urged upon us the notion that language does not represent the world as it is, but instead somehow constructs the world in its purported representation of it.² In the hands of feminists and critical race theorists, this line of thought gives rise to pressing moral and political demands. For if the world that our words (allegedly) construct is an oppressive one, then there are very high stakes in our trying to

² For discussion of this brand of linguistic idealism, see for instance Taylor [1985].

apprehend the process of ‘construction’. We must discern the relationship between speech and social reality not just as a matter of academic curiosity, but because the social reality is one that needs urgently to be reformed.

The key disagreements here, at any rate, are not so much about the *reality* of the ills of identity-based oppression, as about how well we can diagnose the illness by scrutinising racist epithets, sexist jokes, religious hate speech, casually homophobic language, and the like. Linguistically-preoccupied views of oppression may strike some as *prima facie* implausible (and all the more so when they are formulated in broad strokes), but this approach to this topic becomes more compelling when its inchoate conjectures receive a detailed explication, of the sort that we will examine in what follows. At this preliminary stage, though, I will just suggest one way in which we might find ourselves motivated to take up a linguistically-oriented inquiry. The civil rights era of the 1960s and 1970s brought with it massive and unprecedented structural changes in the legal, political, and other social institutions of many Western states. Those who witnessed these changes may have expected that in their wake the pervasive, identity-based social hierarchies which had long disadvantaged women, gays, and non-Caucasians (and others), would soon enough evaporate. So what, then, might we conclude about the causal bases of those hierarchies when *now*, decades after the institutional changes were initiated, and during which time the changes have continued apace, the social hierarchies remain pervasive and potent? If the formal institutional arrangements of our societies cannot, in their reformatory trajectories, eliminate or decisively undermine identity-based social hierarchies, how might we explain the resilience of those hierarchies? Pinning the blame on language is not the only option, but for many theorists it has seemed a sufficiently promising view to warrant attention. After all, language extends into every corner of our social arrangements. If the causal epicentre of identity-based oppression lies in language – in the overt and covert ways in which human difference is coded, reified, and laden with value attributions in our speech practices – then this will go a long way towards explaining why identity-based oppression persists, in the manner and to the extent that it does persist, in the (relatively) institutionally egalitarian world of late capitalist society.

4. McGowan on oppressive speech and conversational score

This is the context in which McGowan’s discussion of oppressive speech appears. The context matters considerably to our understanding the aim of McGowan’s analysis. If our question is simply ‘can speech oppress people?’, then, so one might think, we do not need any clever philosophical machinery to

formulate an answer. Speech can be used to insult, harass, threaten, bully, badger, demean, and humiliate people. If oppression is simply a matter of a person being treated particularly poorly, then of course speech can be used to oppress, since it can be used in any of these ways to treat people very poorly. So this cannot be what McGowan is setting out to show. A better understanding of her purpose is that she is offering us an account of how acts of speech – casual sexist remarks being her main reference case – can be, and indeed are, one of the main, proximate, causal operators through which the oppression of women is effected. Her aim is not just to show how verbal acts (i) make women worse off, or (ii) harm women, or (iii) confound women’s preferences. Rather, her aim is to show how verbal acts might be centrally implicated in the kind of systemic harms and pervasive indignities that we advert to when we say that women are, in societies like ours, in a very general sense, oppressed.

How can Lewis’s work on conversational score-keeping be used to help formulate this kind of account? Firstly, McGowan identifies a species of illocutionary act which she calls ‘the conversational exercitive’; *exercitive* because, in keeping with an Austinian taxonomy of speech acts, the illocutionary type is one which determines what is permissible within a particular domain of conduct; and *conversational* because the domain in which the speech act enacts permissibility facts is the domain of conversation [McGowan 2004: 95]. McGowan appears to endorse Lewis’s account of conversational score-keeping wholesale. She agrees with Lewis that when I, in conversation with you, say ‘Wouldn’t you like to just kick Michael in the shins?’, I bring about an alteration in the informal register of our shared presuppositions and respective expectations, such that certain suppositions are in and out of play, and such that certain subsequent utterances are appropriate or welcome, and others inappropriate or unwelcome. For instance, it would be incorrect play for you to reply with ‘Yes, Michael is great, isn’t he?’. The point that McGowan emphasises, in relation to this phenomenon of conversational pliability, is that my changing the conversational score in this way is a type of exercitive illocution, since, like all exercitives, it is a speech act which determines what sort of actions or behaviours are permissible in a particular domain, in this case, the domain of our conversation. Indeed, in relation to a given conversation, C, potentially *any* utterance can function as a conversational exercitive, enacting permissibility facts by altering, in accordance with Lewisian kinematics, which moves are subsequently apt, licit, or felicitous in C.³

³ McGowan is not the first to apply Lewisian ideas about conversational kinematics to questions concerning speech and sexual hierarchy. Langton and West [1999] use the scorekeeping framework to flesh out the notion that pornography silences women, by changing conversational scores in such a way that makes it difficult for women in turn to have their conversational contributions accommodated in the conversational score. McGowan first engages this material in her [2003] paper, wherein she formulates the concept of the conversational exercitive and recommends it as a useful conceptual tool for refining Langton’s

Next, McGowan argues that conversational exercitives can be used as a means of oppression, again, not in the narrow sense of just treating someone very poorly, but in the wider sense of effecting some type of identity-based structural harm or disadvantage. It is clear enough that standard Austinian exercitives – that is, illocutionary acts which enact permissibility facts – may be used to enact permissibility facts that are oppressive. For example, the autocratic ruler of country X may authoritatively decree that women must always be accompanied by a male family member in public. And it moreover seems clear enough that such an exercitive would effect identity-based oppression, rather than mere harm or mere disadvantage, since the permissibility facts which it enacts would generate systemic and pervasive disadvantages for the women of X (and if you like, it would impose these disadvantages on X’s women ‘as women’). McGowan’s contention, then, is that just as these standard Austinian exercitives can oppress, so too can conversational exercitives oppress, if and when the permissibility facts which they enact, in the local context, are oppressive in the appropriate sense.

The pressing question at this point, though, would be: how is an enaction of permissibility facts within the relatively narrow domain of a conversation tantamount to, or an instance of, the kind of structural and far-reaching oppression that McGowan purports to diagnose? After all, it would be a somewhat trivial finding – and again, an insight that we could arrive at without using any philosophical gadgetry – that speech can ‘oppress person A’ in the very limited sense of ‘spoiling a conversation for A’ or ‘making certain things un-say-able for A within a given conversation’.

McGowan’s response to this concern is to offer a more finely articulated account of the relationship between, on one hand, systems or structures of oppression, and on the other, the points at which the operations of such systems get a purchase on their subjects. Perhaps it isn’t obvious why this relationship stands in need of any fine articulation. It’s true that we could simply draw a distinction like the one Sally Haslanger [2004] draws, between *individual* oppression, i.e. person A oppressing person B, and *structural* oppression, i.e. person B being oppressed by wider patterns of social activity, which are beyond the control of any individuals, including A.⁴ In its elementary form, however, that kind of

[1993, 1998] speech-act-theoretic analyses of MacKinnon’s claims about pornography’s subordination of women. McGowan returns to the subject in [2004], outlining a more general-purpose account of the conversational exercitive, and observing some of the key features which distinguish it from regular Austinian exercitives (e.g. the fact that it does not function via the recognition of speaker intentions, and how this explains the unusual felicity conditions for its performance). McGowan’s extended examination of how conversational exercitives function in oppressive speech comes in [2009].

⁴ Haslanger says we should not principally conceive of oppression as a relation between individuals, because this ignores the fact that oppression is often effected by complex systems that are beyond the control of individual agents. But at the same time she

distinction obscures our understanding of the interactions between different individuals living together in circumstances of structural oppression. If B is subjected to structural oppression, then it is true by definition that A cannot by himself act so as to cause or prevent B's oppression. But A can still be a participant – and a morally culpable participant – in the realisation of the patterns of conduct through which B's structural oppression *is* effected. The fact that A's participation is redundant with respect to the system's functioning does not mean that A's behaviour is not, after all, participation.

The question, then, is how we should conceive of the different kinds of participations or positive contributions that individuals make to structural oppression, and how we may do this without losing sight of 'the structural' (e.g. by simply describing some unusually subtle modes of individual-on-individual oppression). McGowan approaches this question through a simple but telling example. John and Steve, working together at a factory with few female co-workers, have the following exchange.

John: So, Steve, how'd it go last night?

Steve: I banged the bitch.

John: [smiling] She got a sistuh? [McGowan 2009: 399]⁵

McGowan then says that

Steve's utterance makes it acceptable, in this immediate environment and at this time, to degrade women. His utterance makes women second-class citizens (locally and for the time being). If Steve's utterance does this, then it is akin to a sign reading: 'It is hereby permissible, in this local environment and at this time, to treat women as second class citizens.' Such a sign would surely be an act of gender oppression. The hypothesis suggested here is that, perhaps, Steve's utterance is too. [McGowan 2009: 400]

wants to acknowledge the genuine possibility of individual-on-individual oppression, and this is why she draws the distinction adverted to here [Haslanger 2004: 104-05].

⁵ It is perhaps unfortunate that McGowan calls this a working class example and uses the unusual phonetic spelling of 'sistuh' rather than 'sister' to drive the idea home, on both counts leaving herself open to a charge of class prejudice. She deflects this by including a footnote that describes a parallel scenario involving two stereotypically posh men conversing at an elite gentleman's club [note 23, at 105]. I wonder whether, in the interests of avoiding the appearance of classism (an even more important desideratum than usual, in a paper about the relationship between language and structural oppression), it may have been better to have used the 'posh' example rather than the 'working class' example in the main text.

There are a few points about this suggestion that need to be teased out. McGowan does not want to say that Steve is responsible, in a single remark, for creating the vast, temporally-extended system of sexist objectification. The structural oppression of women, on McGowan's account, is a rule-governed activity in the sense that it is a complex, coordinated system of social interaction, which ranks people and imposes norms, with associated expectations and informal penalties, for how people are to be treated in view of their ranking [McGowan 2009: 397]. McGowan further suggests that the oppression of women, like other rule-governed activities, is governed by *two types* of rules: (i) general rules, or *g-rules*, which 'govern all instances of the rule-governed activity in question and... are not enacted by the performing of any particular such activity'; and (ii) specific rules, or *s-rules*, which are 'enacted by the performing of the very rule-governed activity over which they preside', and which are limited in scope and duration [McGowan 2009: 396]. The g-rules establish broad parameters for how the activity in question is to proceed, and the s-rules determine, within those parameters, what may happen (at any particular) 'here' and 'now'. The g-rules governing the structural oppression of women may, as McGowan concedes, be difficult to identify with great precision. In their basic orientation they will be rules which prioritise certain of men's interests over women's interests, and which confer certain prerogatives on men. In any case, what McGowan wants to say of Steve's remark is that it enacts s-rules, in the local context, to the effect that women may be treated as inferiors 'here and now'. It is true, on her view, that Steve's remark can only enact these s-rules because its content is *consonant* with the g-rules that fix the broader parameters of the gender system in the society that Steve and his co-workers inhabit. Nevertheless, what Steve's remark does is bring the latent force of that gender system to bear in the local context. The oppressive power of Steve's comment is derived from something for which he is not responsible – the underlying systems of social organisation – but the activation of that oppressive power in the local context *is* something for which Steve is causally responsible and (defeasibly) morally responsible. McGowan's contention, in summary, is that Steve's speech functions as a conduit through which the underlying structural oppression of women gains a purchase on particular individuals, both men and women, at a particular place and time.

5. The continuity of verbal and non-verbal moves in oppressive systems

A further point that needs to be spelled out is how, within this account of oppression as a rule-governed activity, we can make the leap from a comment being able to alter the *conversational* score, to the state of affairs McGowan describes, in which a comment changes not just what it is permissible to say about

women here and now (e.g. to carry on speaking in the same degrading tone as Steve), but also what it is permissible to do *beyond the realm of speech* – for instance, to treat women in an economically subordinate manner. Lewis’s observation, recall, was just that conversational scores are specially pliable. It does not follow that all parts of a social interaction in which a conversation occurs are pliable in the same way as a conversation itself is. McGowan’s suggestion on this front is that in a system of identity-based oppression, there is a continuity between verbal acts and non-verbal acts; all of them are, she wants say, mutually-reinforcing moves in the rule-governed activity of women’s oppression. Indeed, the conversational exercitive, for McGowan, is just one form of the more general phenomenon of covert exercitiveness, in which permissibility facts are enacted without an explicit signalling of their contents. Something being a move in a rule-governed activity, on McGowan’s view

requires only that the action in question be a contribution to, and thus a component of, the activity in question. Since speech is one way to differentially treat people, and since it is one way to differentially treat people in virtue of a person’s membership in a socially marked group, speech is certainly sometimes a move (in the rule-governed activity) of oppression. Racist and sexist speech, for example, are such moves. [McGowan 2009: 397]

In light of this conception of gender oppression as a rule-governed activity, the idea that speech acts can oppress women generally – as I put it above, that they can oppress ‘beyond the realm of speech’ – follows simply from speech being characterised as a move within the rule-governed activity. The g-rules of gender oppression disadvantage women in many different ways. Conversational exercitives are one way to enact s-rules which specify, within the parameters of those broad-spectrum g-rules, how women may be differentially treated in specific contexts. McGowan does not see speech as the *only* way to perform moves in the rule-governed activity of women’s oppression. But her account of conversational exercitives is meant to show how speech can at least *be* a move within that activity.

A potential objection is looming here, however. Granted, speech is, as McGowan puts it ‘one way to differentially treat people’. But differential forms of treatment are not all alike. If A invites B but not C to join a basketball team, that is ‘a way of differentially treating’ C; if A’s non-invitation owes to the fact that C is a woman and that the game is being held in a men’s sporting league, then that is a way of differentially treating C ‘in virtue of her membership in a socially marked group’. But it does not immediately follow that A has performed a move in the rule-governed system of women’s oppression. Maybe he has, maybe he hasn’t; the answer will depend on whether gendered sporting leagues are significantly implicated in the complex causal web of women’s oppression, and the view that they *are* is

something that needs to be argued for, not just asserted. Now, it is more plausible to say that sexist speech is significantly implicated in the causal web of women's oppression, but the claim is contestable, and indeed contested. There is a view of verbal conduct – evident in much jurisprudential writing about freedom of speech – which adheres to the old playground rhyme that begins 'sticks and stones may break my bones'.⁶ According to proponents of this view, we can best explain the oppression of 'socially marked' groups by adverting to inequitable material aspects of our social arrangements, such as the inequitable structure of the employment sector; speech, language, and discourse are either peripheral causal factors or just downstream consequences of these other root causes.⁷ Part of McGowan's purpose is to contest this view: to show that far from being peripheral or inert, speech is one of the principal causal levers through which oppression is enacted. But at this point in McGowan's argument it seems that the idea that speech is a covert move in the rule-governed activity of women's oppression is being invoked as a founding conceptual claim on which other parts of the account are supposed to rest. I do not think the notion is implausible, but given the preponderance in certain influential circles of the 'sticks and stones' view of speech, it seems unsatisfying here to just insist that speech *is* a move in the rule-governed system of women's oppression, especially if empirical data could be cited to support that view of things. What sort

⁶ Here's one revealing example of this view in action. The American legal scholar Cass Sunstein [1996] defends an expressivist conception of law, on which it is an important and proper part of the law's function to express and endorse certain ideals and values to its subjects. Expressivist aims, for Sunstein, are a key part of the justification for legal restrictions on things like racist hate speech [see also Post 2009: 129]. But why emphasise expressivist aims? Why not advert to the harms or other adverse consequences of hate speech? For Sunstein, debates about the legitimacy of anti-hate speech law 'could not plausibly be focused on consequences', because, he thinks, *'the stakes are relatively low'*, and cannot justify the amount of time and energy devoted to the issue' [Sunstein 1996: 2023, my emphasis]. This is a remarkable claim, because the most influential texts arguing in favour of legal restrictions on hate speech [e.g. Delgado 1993; Lawrence 1993; Matsuda 1993; Delgado and Stefancic 2004] are *all* focused, in arguing the case for anti-hate speech law, on the harmful consequences that may be attributed to hate speech. Sunstein's claim that 'the stakes are relatively low' does not follow an argument to the effect that these other authors are mistaken in attributing egregious harms to hate speech. It is stated, rather, as though it were simply a commonsense observation. But why would this claim seem like a matter of assertible commonsense to Sunstein and his audience of legal scholars, notwithstanding evidence and argument to the contrary? In short, it is for the reason stated above, namely, the fact that the 'sticks and stones' view of speech operates as the default position in American jurisprudential scholarship.

⁷ This deflationary view concerning the role of verbal factors in oppression is sometimes expressed as a complaint, roughly, that linguistic concerns distract from the real work of combating social injustice. See for instance Marjorie Heins's suggestion that 'too much work remains in the battle against the subtle and unspoken... institutional racism entrenched in our society to squander resources suing loudmouths' [Heins 1983: 592], or Henry Louis Gates's claim that attempts to curb racist speech merely pay lip service to civil rights, since they don't address the pressing realities of economic disadvantage [Gates 1993: 43].

of data might do the job? Suppose we agree that the imposition of discriminatory labour conditions is a form of identity-based oppression, e.g. men being rewarded and women being penalised for aggression in the workplace. Now, suppose also that we have data which indicate that subjects' prior exposure to identity-prejudicial language, e.g. their hearing women being referred to as 'bitches', has a significant influence on the likelihood of subjects making negative judgements about women who demonstrate social confidence. Such data would indicate that verbal factors are continuous with certain paradigmatic 'moves' in the rule-governed activity of women's oppression, where by 'continuous with', I mean the verbal acts causally interact with the non-verbal acts in a non-epiphenomenal way. Such data would, therefore, go some way towards establishing that sexist speech really is a move in the rule-governed activity of women's oppression – and not just that this is conceivable or plausible, but that it is something advocates of the 'sticks and stones' view must accept on pain of disregarding the evidence.⁸

6. Speaker authority

McGowan claims that speech oppresses by enacting permissibility facts, which is to say it determines, within a broader, identity-oppressive system of practices, conventions, and expectations, which courses of action are apt or permissible in particular context. In a complex mesh of identity-oppressive social arrangements, speech acts provide cues which permit individuals to say and do oppressive things to members of socially marked groups. When person A utters a sexist remark in the workplace, not only

⁸ Part of what is at stake, in this question, is whether it is only speech in the conversational *mode* that has the oppressive function McGowan identifies, or whether any speech – conversational or not – can enact oppressive permissibility facts. Obviously there is much speech that does not take place in the conversational mode; there are books, speeches, articles, lectures, editorials, and many artistic forms of speech (poems, plays, and novels), which are addressed to their audiences primarily in a 'one-way' fashion. Why might this matter? Because to the extent that conversations are specially pliable in the way Lewis observed, this is not an accidental quirk, but rather an upshot of the *cooperative character* of conversational speech. What makes a locutionary exchange a conversation, as opposed to a dysfunctional quarrel or a free-associative talking game, is the joint aim of the participants to be talking about the same things, and to be talking with similar communicative ends in mind (e.g. solving a problem, or exchanging information). It is because of this cooperative character that unchallenged presuppositions can presumptively acquire the status of common ground for participants in a conversation. By contrast, in forms of verbal interaction that are not inherently cooperative, demurrals and objections do not need to be issued immediately on pain of acquiescing to an interlocutor's viewpoint. At certain points in McGowan's account it is unclear which parts of our verbal practice she thinks may be properly characterised as conversational, in a way that would allow her insights about conversational exercitives to apply.

does he make it appropriate in what follows to *speak* degradingly of women, he makes it acceptable per se, in what follows in that immediate environment, to degrade women [McGowan 2009: 400].

In this section I want to draw attention to the unique perspective on the issue of ‘speaker authority’ that is available under McGowan’s account of oppressive speech. This is a notable strength of her view, but it is also, as I will explain, a potential focal point for her critics.

To broaden the context a little, McGowan’s work is located within an intersecting body of scholarship in philosophy, jurisprudence, and political theory, which seeks to defend legal restrictions on different kinds of hostile expressive activity (pornography and racist hate speech being the two most common focal points) by explicating the nature of the relationship between these hostile expressive activities and the harms or disadvantages experienced by the groups (e.g. women and ethnic groups respectively) against whom the hostility of these expressive activities is directed. McGowan is quite cautious about whether a legislative agenda might follow from her conclusions on the nature of oppressive speech.⁹ But for those authors who are broadly allied with McGowan’s views about the significance of speech in oppression, and who *do* endorse legal restrictions on pornography or hate speech, the characterisation of speaker authority is a recurrent problem. Feminist critics of pornography, for instance (in the mould of MacKinnon, Langton, et al.) say that pornography – in its regular, everyday use as a private sexual aid – subordinates and silences women. And yet pornography, as Leslie Green [1998] reminds us, is tawdry and low-grade speech, which, although legally tolerated, is still widely regarded (including by many of its users) as tawdry and low-grade. How can it be, then, that pornographers have sufficient authority, or power, or influence, to enact oppressive systems of social organisation? Moreover, how is this possible when there are other sources of sexual norms – e.g. in the reformatory trajectories of the political and legal systems, in religious institutions, and with lesser influence, in the academy – persuasively contesting the crassly objectifying view of women that pornography encodes? Similar questions can be raised about racist hate speech. Neo-Nazis and other far-right political groups are widely seen as ignorant, overheated bigots. When these individuals begin to approach mainstream prominence, their successful transition into the mainstream is dependent upon their ability to back away from the fervid language of hate speech and affect the more restrained tone of

⁹ In the paper I am centrally concerned with here, McGowan does not explicitly endorse the restriction of oppressive speech, but nor, she says, is she ‘yet convinced’ that regulation would be ‘completely wrongheaded’ [McGowan 2009: 401]. In other work [Maitra and McGowan 2007] McGowan does conditionally defend legal restrictions on pornography, but this more general policy position stands or falls independently of her account of oppressive speech in [2009].

‘serious political dialogue’. But how does this rabble of extremists – those we may call ‘hate speakers’ – have the power or authority to perpetuate society-wide systems of social hierarchy, as critical race theorists standardly claim, or to deprive their targets of their dignitarian status as equal members of society, as authors like Steven Heyman [2008] and Jeremy Waldron [2010] claim? And again, how can this be the case when there are prominent social institutions – in law, politics, religion, the media, and education – forcefully opposing the crude, demeaning ideas about ethnicity and human worth that constitute the unifying core of racist hate speech?¹⁰

Answers to these sorts of questions are sometimes given, some more persuasive than others.¹¹ The usual approach is to tell a story on which the pornographer/hate speaker speaks with the force of an oppressive structure behind him [Schwartzman 2002: 431]. The claim is not that lone pornographers are responsible for creating the whole sprawling edifice of sexual oppression, nor that hate speakers have authority in the system of structural oppression under which ethnic groups are subordinated, in anything like the way that the Archbishop of Canterbury, say, has authority within the Church of England. Rather, the claim is that social structures are the source of oppressive power, and that this is channelled *through* pornographers/hate speakers, so that they are properly *answerable* for the harm that eventuates, even though they do not themselves originate the power underlying the harm. This is only a brief schematic outline, but it is enough for us to see that what critics of pornography and hate speech are looking for is just the kind of analysis that McGowan delivers. These critics want to explain how speakers mobilise the oppressive power of social arrangements, so that speakers may be regarded as agentially responsible for acting oppressively, despite the fact that speakers are *not* responsible for the existence of the social

¹⁰ It would be wrong to suggest that pornographers and hate speakers have *no* power, authority, or influence, in the informal senses of these words. There are some prominent pornographers and wealthy hate speakers, after all. My point concerns the inflated status that is attributed to these figures when their speech is deemed responsible for inflicting society-wide inequalities. Scholars cannot speak from their university chairs and transform the world to fit their vision of how it should be. Political leaders cannot shape societies at will; they have to eke out compromises and settle for incremental policy changes. Pornographers and hate speakers seem to have less moral and intellectual authority than scholars, and less social influence than politicians, and yet, so we are told, they can shape whole societies to fit their vision of how the world ought to be. This is the extraordinary claim that we find in critiques of hate speech and pornography which stands in need of explanation.

¹¹ I think it is easier to be persuaded by Jeremy Waldron’s [2010] approach to this question (focused upon hate speech), than by Rae Langton’s [1998] approach (focused upon pornography). Waldron tries to explain how hate speakers can have a profound impact upon the civic status of their targets even though they *do not* occupy a position of any power or authority. Langton, by contrast, argues that despite appearances pornographers in fact *do* have a genuine kind of authority, in a limited domain, which is nevertheless commensurate with the great degree of social influence she attributes to them.

arrangements whose oppressive power they mobilise. McGowan's conceptual tools – the conversational exercitive and her account of the continuity between linguistic and extra-linguistic factors in structural oppression – allow for an explanation of the sought-after type. McGowan rightly identifies this as an advantage of her view over other accounts of the relationship between speech and identity-based oppression.

Because the system of gender oppression is just a complex human activity, everyone whose activity constitutes (part of) that system is clearly capable of making moves in this activity. This is clearly true of Steve... Since Steve is a participant in the rule-governed activity of gender oppression, his utterance is able to change what is subsequently permissible. Thus, although he has no peculiar authority over the system of gender oppression, his utterance is nevertheless able to enact permissibility facts (s-rules) in it. [McGowan 2009: 402]¹²

But then, if this is a strength of McGowan's account, why might it also represent a potential point of weakness? The difficulty, as McGowan is well aware, is that if no special authority is required to activate identity-oppressive presuppositions and expectations in a local context, owing to the specially pliable character of conversation, then correspondingly, no authority should be needed to remove those presuppositions and expectations, and thus forestall any oppressive outcomes they may precipitate. In short, the special pliability of the conversational score cuts both ways.

McGowan's response to this worry, inspired by Lewis, is to posit an asymmetry in the pliability of conversational scores in cases in which identity-oppressive speech is operative. It is true that no special authority or power is required for racists or sexists to contaminate the communicative environment for their targets, and this is true for the reasons that Lewis highlights, regarding the special malleability of conversational activity. However, McGowan says, *riding* the communicative environment of identity-prejudicial presuppositions and expectations is very difficult. 'Setting the record straight in response to a

¹² McGowan sees her account as supplementing feminist critiques of pornography, and I am ostensibly endorsing that view of her work in what I say here. Langton, for one, has recently written approvingly of McGowan's approach to these issues, noting in particular the fact that McGowan's approach is well-equipped to explain how non-authoritative speech nevertheless can have far-reaching social consequences [Langton 2012: 137-38]. Notice, however, that it would be open to McGowan to question the anti-pornography feminists' preoccupation with the sexualised extremes of misogynistic expression. If McGowan is right that the Steves of this world are enacting the oppression of women in their 'everyday', conversationally-exercitive speech, she might well regard it as something of a red herring to home in, as many authors have, on the distinctive operations of pornography. Thanks to Aweek Bhattacharya for bringing this point to my attention.

sexist remark' she says, is 'akin to trying to unring a bell... there is something complicated and covert going on that is difficult to pinpoint and hence undo' [McGowan 2009: 403].¹³

7. Varieties of conversational asymmetry

The asymmetries McGowan posits seem to have a different structure to those Lewis identified. His examples all involve stringency in the application of terms. Person A: 'Australia is flat'; B: 'But it's not *really* flat, not like this polished steel is flat'. In a conversation like this, one interlocutor can easily shift the conversational score so that relatively lax uses of 'flat' become incorrect or inapt. But reversing this shift in the score is much more difficult. Why the asymmetry? Here's a rough hypothesis that accords with arguments Lewis makes elsewhere [e.g. what he says about knowledge in Lewis 1996]. In ordinary discourse we use words like 'flat' with as much or as little stringency as is needed given what we are trying to do with, and what is at stake in, our communication (e.g. the sharing of information). However, if the relativistic character of a particular term is made salient – for example, by one of us verbally drawing attention to its relativity – we will then favour more stringent usages, because one of the things that will (newly) be at stake, in our communication, is the consistent application of a relative term. And we will continue to favour stringent usages of the relative term so as long as its relativistic character remains a salient consideration (salient just in a subjective sense, i.e. salient *for us*). The asymmetry is due, then, to the fact that there is nothing either of us can say that will make the relativity of a term like 'flat' become *un*-salient, as swiftly or as easily as it became salient. One may argue that always and automatically favouring stringency in the application of these terms is a mistake, as Lewis sometimes does. But in doing so one perpetuates the salience of the fact that these *are* relativistic terms, and thus one sustains our reflex judgements about what would count as sufficiently stringent usages of the terms. In order for the relativistic character of a term that we're using to become *un*-salient, we will simply have to change the

¹³ McGowan hedges a little by suggesting that even if an act of oppressive speech *can* be easily reversed, this 'does not entirely disqualify it as oppressive' [2009: 403]. She imagines a case in which a business owner hangs a 'whites only' sign, but in which the policy enacted by the sign can be revoked by anyone removing the sign (the example assumes that the sign's removal can be easily carried out). The policy in such a case may be short-lived, McGowan says, but it is oppressive nonetheless [2009: 403-04]. This is unconvincing. If the phenomenon of 'oppression' that we're trying to diagnose is stable, operant over extended periods, consistent in its character – features of the sort that McGowan herself stresses – then it is unclear how ephemeral policies can be understood as proper parts of the phenomenon, or indeed, how they can even be seen as policies, as opposed to momentary states of affairs.

subject, or distract ourselves, or allow some time to elapse, so that we come to the term, as we might say, with ‘fresh ears’.

The asymmetries that concern McGowan, by comparison, apparently have nothing to do with stringency in the application of relative terms. Her claim is that when person A makes (e.g.) a casual sexist remark, he introduces into the conversation an extensive and complex schema of sexist ideas and attitudes, including stereotypes, hostile feelings, aesthetic judgements, false beliefs, and distorted perceptions. The introduction of this schema makes it part of the common ground for the conversation and social interaction in which person A is participating. Statements, actions, and attitudes become appropriate, correct, or eligible for other participants to the extent that they accord with the sexist schema of thought. In other words, the sexist remark shifts the conversational score, swiftly and easily, along a sexist trajectory. However, so the key claim goes, it is particularly difficult to reverse this shift in the conversational score, so that the sexist schema’s freshly-operative presuppositions and expectations are removed from being part of the common ground for participants in the social interaction.¹⁴

How plausible is this hypothesis? And if it is plausible enough to merit serious consideration, how might we try to explain the alleged asymmetry? There are data in the social psychology literature that provide some support for McGowan’s view. Although an asymmetry in the activation and deactivation of negative schemas has not been investigated in any study that I’m aware of, there is ample evidence to suggest that negative schemas in accordance with typical identity-prejudicial attitudes can be activated – by which I mean, made to influence people’s judgements – with remarkably *little* effort. In one kind of study, one-off, isolated uses of derogatory slurs or allusions to negative stereotypes have been found to lead people to downgrade their assessments of or expectations towards members of targeted groups. This kind of effect has been demonstrated both for third parties [Greenberg and Pyszczynski 1985; Goodman et al. 2008], and for members of targeted groups themselves [Steele and Aronson 1995]. It remains to be shown, but it seems *prima facie* unlikely that identity-prejudicial schemas could be disabled quite as effortlessly as they are, so the evidence seems to indicate, activated.¹⁵

¹⁴ It is not only verbal activities in which the phenomenon of asymmetric pliability may arise. In one of her earlier pieces, McGowan observes that ‘a similar phenomenon seems operative in the heterosexual sociosexual arena. Once some formerly taboo sexual practice is introduced and treated as permissible, it thereby is permissible in virtue of the rules of accommodation operative and, once this happens, it is difficult to subsequently re-introduce the taboo’ [McGowan 2003: 188].

¹⁵ I mention just a few studies here, but note that there is a rapidly-expanding body of research on the psychological and sociological impact of identity-oppressive language. Some studies skirt around the larger causal questions by exploring how the experience and perception of hate speech differs across lines of social difference and circumstance. There are qualitative studies

If there is empirical research that can be adduced in support of McGowan's asymmetry hypothesis, one might wonder whether theoretical reflection can shed any additional light on the topic. I think it can. Suppose we have compelling and convergent evidence confirming the kind of asymmetry that McGowan postulates. What should we then say about the phenomenon? It seems unsatisfying to just conceive of this as a brute, mysterious fact about our cognitive habits. Of course we could say that there are asymmetries of this sort *because* societies are sexist or racist. But surely we don't want to leave it at that? This conclusion seems to treat identity prejudice as a cosmic mystery or an inexplicable force of (human) nature. What we want is some insight into what is involved in a society's being sexist. Should this be understood as a claim about the kinds of beliefs and attitudes that are held by a majority of a given society's members? Is it a claim about the policies and practices that operate in a given society, and the kinds of beliefs and attitudes which – if they were attributable to a unified, governing agency – would render those policies and practices intelligible? Or is the claim getting at something else altogether? Instead of saying that conversations involving identity-oppressive speech are asymmetrically pliable *because* societies are sexist, we can try to forge a clearer understanding of what a society's being sexist consists in, via an understanding of why the asymmetric pliability might exist in these cases.

To this end, I suggest we think about the phenomenon of asymmetrically pliable scores in oppressive speech as being primarily a matter of *associations*. I'll begin with an example. A celebrity blogger reports that a certain pop singer, LG, is transsexual. The report is quickly discredited, so that we return to our earlier evidentiary state, of having no *good* reason to suppose that LG is transsexual. And yet the association remains. It remains, in part, because the discrediting of the association subtly perpetuates the association. Statements like 'LG is really a woman', or 'reports that LG is transsexual have been discredited' reinforce the notion that, on some level, LG and transsexuality have something to do with each other. 'No smoke without fire', so the malign thought may go, even when the 'smoke' is something

which indicate that members of disadvantaged social groups are more likely than others to feel personally attacked in public verbal confrontations [Nielsen 2004], and studies in social psychology which show that subjects' judgements about the offensiveness of hate speech vary depending on gender, race, sexual preference, and that these judgements are sensitive to a variety of framing effects [Cowan and Mettrick 2002; Cowan and Khatchadourian 2003; Cowan et al. 2005]. There are other studies which aim more at the heart of the causal complex of language and social disadvantage, by identifying correlations between identity-prejudicial speech and various markers of structural oppression (e.g. suicide rates, rates of naturalisation and inter-marriage, residential segregation), and by seeking to ascertain the extent to which these correlations are sensitive to the character of (e.g. the degree of complexity in) the identity-prejudicial language in question [Mullen and Rice 2003; Mullen and Smyth 2004; Leader et al. 2009].

the evidentiary force of which has, by one's own lights, been discredited. An association that is recognised as having spurious or sinister origins can still remain an association.

Carried forward, my suggestion is that the identity-prejudicial schemas of thought which, on McGowan's view, can easily be brought to bear on our social interactions, might usefully be conceived of as associative schemas. The presuppositions and expectations are all, so we might claim, generated by a web of associations between (for example) women and various negative traits and qualities. Women, so the schema has it, are to be associated with sexual passivity, weakness, domesticity, etc. The sexist remark brings forth these associations, and exercitively generates presuppositions and expectations that accord with them. And it does so very easily, (i) because we are all already familiar with the associations in question from their suffusion of popular culture, educational practices, etc., and (ii) because conversational scores, quite generally, are always ready to be reshaped by fresh moves and the exercitive enactments that come with them. Subsequently, however, these associations, even if we recognise them as spurious or sinister, are resistant to our efforts to remove them from our minds, or quarantine their distorting influence upon our judgements. And this is because comments which aim to discredit an association tend to accidentally reinforce it. Given the nature of associative thinking, someone's repudiation of the idea that women exist for men's sexual gratification can still give some small credence to the notion that there *is* an association to be made between women and sexual objectification. If this is right, then McGowan's asymmetry hypothesis is understandable in light of the more general asymmetry in creating and eliminating associations. It is easier to shift conversational scores along sexist trajectories because sexist presuppositions and expectations are governed by invidious associations, and associations in general are relatively easy to initiate through verbal activity, and relatively difficult to eliminate via verbal activity, if indeed they can be eliminated in that manner at all.¹⁶

¹⁶ In these conjectures about associative schemas I am opening up a point of intersection between McGowan's work and the growing body of research on the nature and ethics of implicit bias [e.g. Blum 2004; Kelly and Roedder 2008]. One of the concerns in this literature is how we should characterise the processes through which implicitly-biased judgements are rendered, and accounts that ascribe a major role to associative schemas represent one candidate view about this matter. In the present discussion I do not address the question of how exactly we should conceive of associations between A and B, or associative schemas linking A with B, C, D, etc., beyond the definitional idea that 'associanda' must be reliably correlated. Nor have I said much about the actual content of identity-oppressive schemas, and how the contents of such schemas vary from one 'socially marked group' to another. Of course, these issues are the focus of much feminist theory, critical race theory, queer theory, and other critical discourses, and if a view like McGowan's is going to be elaborated using the idea of associative schemas, it will be more compelling to the extent that it can unite a formal analysis of how conversational exercitives operate, with substantive analyses of the content of identity-oppressive schemas of thought.

This is all in keeping with the picture that McGowan paints. She says, recall, that responding to a sexist remark is like trying to unring a bell. This is an evocative metaphor, but we can put a finer point on the insight it contains by reflecting upon the character of associative thinking and conversational salience. One thing we find when we do so is that there are similarities after all between Lewis's examples of conversational pliability, involving stringency, and the cases that interest McGowan. In both cases, the asymmetric pliability of conversation is due, so I suggest, to the asymmetry between what it takes to make something salient in speech, and what it takes to make the same thing (subsequently) un-salient.¹⁷ In a conversation there are many different kinds of thoughts – including things like the recognition of a modally relative term, or the apprehension of a complex associative schema – which can be made salient in an instant, but which cannot be made un-salient in any quick and straightforward manner. The salience of such considerations will – in accordance with Lewisian rules of conversational kinematics – alter how it is permissible, apt, or felicitous for participants to behave in a given conversation. In Lewis's pet cases the result is (for instance) that non-stringent usages of certain words will be rendered 'out of play'. In McGowan's cases, if she is right about the continuity between verbal and non-verbal moves in the social practices of identity-based oppression, the result is that people are oppressed.¹⁸

8. The bigger picture

The broader agenda that is pursued in McGowan's work, as I mentioned in §1 and §5, is to show why *speech matters* in our efforts to explain and combat identity-based social hierarchies and subordination. And in this connection there is a question that remains open, even if we accept the bulk of McGowan's account, about whether the linguistic phenomena she describes *produce* identity-based oppression, or

¹⁷ On this front as well, my elaboration of McGowan's account chimes with her own discussion. She mentions at various points, for instance, that changes in conversational score can take the form of changes in facts about what is salient to the conversation's participants, although in this she emphasises narrower aspects of salience (e.g. who is picked out by the use of a pronoun) than the associative schemas that I'm adverting to here.

¹⁸ One might perceive a tension between my remarks about salience and certain things McGowan says about the covertness of oppressive speech. Conversational exercitives almost always operate covertly, she says, in that the permissibility facts they enact are not explicitly signalled in the locutionary content of the speech act. But is it possible for something to be made salient in a covert manner? On McGowan's usage of 'covert', it is. Steve's remark makes a negative associative schema salient and this in turn enacts permissibility facts. However, the exercitive force of Steve's utterance is still covert, because the utterance does not explicitly convey, in its locutionary substance, the content of the permissibility facts that are enacted in its being uttered.

whether they are generic linguistic phenomena that are present in circumstances of identity-oppression merely because they are present in all (or most) verbal interaction. Consider the following example, adapted from Richard [2008]. Suppose the Mantown Philosophy Department hires two women in 2012, and the Vice-Chancellor of Mantown University later says to a colleague: ‘this year Philosophy hired *three broads*’. Richard says that despite the numerical error an utterance like this should not be adjudged *false*, because an assertion of its falsity assents to the same contemptuous misrepresentation of women as the utterance itself. We can say that statement is *erroneous*, but its erroneousness is not the truth-apt erroneousness of being false [Richard 2008: 7]. Now, the linguistic phenomenon in the example – the representation of Xs *as* scorn-worthy via the erroneous (but not *false*) verbal exhibition *of* scorn for Xs – obviously need not be confined to remarks about women. For instance, the Vice-Chancellor of Rotund University might say ‘this year History hired three bean-poles’. In this case it seems unlikely that identity-oppression is occurring, because slim people are not usually oppressed on account of their slimness. However, the linguistic phenomenon that facilitated the oppressive speech in the first case occurs in the second case as well, and thus it is unclear whether that linguistic phenomenon can really *explain* the fact that an identity-oppressive social transaction occurred in the Mantown Vice-Chancellor’s remarks, any more than the Vice-Chancellor’s competence as an English speaker can explain that putative fact. If we want to diagnose the identity-oppression that is going on in a case like this, looking to the linguistic phenomena will only get us so far, and probably not as far as we want to go.¹⁹

Parallel doubts may be raised about McGowan’s account. Conversational exercitives and the phenomenon of asymmetric pliability of conversations may be collocated with the oppression of social groups (women and others), and they may facilitate it. Whatever else is true, then, we do well to understand how these phenomena function. However, unless we have reason to think that the operation of these phenomena is confined to the use of identity-prejudicial language against members of particular out-groups, such observations do not by themselves vindicate a language-oriented approach to combating identity-based oppression. Granted, we can always say that speech is ‘involved’ in identity-based oppression, but this just postpones the problem, which is now whether that ‘involvement’ consists merely in collocation with oppression, or in the facilitation of oppression, or whether it is something more causally essential to the enactment of oppression. Such questions are not merely academic. Modern liberal democracies have seen a steady accretion of legal, institutional, and informal practices (e.g. anti-hate speech laws, work-place codes of conduct, political correctness norms), the aim of which, at a face value

¹⁹ Thanks to Christopher Jay for pressing me on this point.

interpretation, is to effect some sort of social change by regulating people's language. If we are going to carry on along this trajectory, it is crucial that we properly understand what we are doing. Are we trying to control the *symptoms* of identity-oppression when we regulate identity-prejudicial speech, or are we treating the root of the illness? And (to extend the metaphor) if it is the latter, what is the causal process through which the illness inflicts pain and damage upon its sufferers? McGowan's analysis speaks mostly to the second question, and it does an especially good job of illuminating the relationship between individual acts of speech and broader systems of structural oppression. The question that remains, so I suggest, is whether the underlying structures of identity-based oppression can still survive, and still exert their toxic influence, if we manage to reform the speech practices through which they get their purchase on subjects, from one time and place to the next. Nevertheless, what McGowan has given us, and what I have tried to bolster – in my remarks about salience and associative schemas in relation to the asymmetric pliability issue – is a plausible and sophisticated account of the process through which that purchase is gained.²⁰

References

- Austin, J. L. 1962 (1976). *How to Do Things with Words* (2nd ed.), Oxford: Oxford University Press.
- Blum, Lawrence 2004. Stereotypes and Stereotyping: A Moral Analysis, *Philosophical Papers* 33/3: 251-89.
- Cowan, Gloria, Becky Heiple, Carolyn Marquez, Desiree Khatchadourian, and Michelle McNevin 2005. Heterosexuals' Attitudes Towards Hate Crimes and Hate Speech against Gays and Lesbians: Old-Fashioned and Modern Heterosexism, *Journal of Homosexuality* 49/2: 67-82.
- Cowan, Gloria and Desiree Khatchadourian 2003. Empathy, Ways of Knowing, and Interdependence as Mediators of Gender Differences in Attitudes Towards Hate Speech and Freedom of Speech, *Psychology of Women Quarterly* 27/4: 300-08.
- Cowan, Gloria and Jon Mettrick 2002. The Effects of Target Variables and Setting on the Perception of Hate Speech, *Journal of Applied Social Psychology* 32/2: 277-99.
- Delgado, Richard 1993. Words That Wound: A Tort Action for Racial Insults, Epithets, and Name Calling, in *Words That Wound: Critical Race Theory, Assaultive Speech, and the First Amendment*, Boulder Colorado: Westview: 89-110.
- Delgado, Richard and Jean Stefancic 2004. *Understanding Words That Wound*, Boulder Colorado: Westview.
- Dworkin, Andrea 1981. *Pornography: Men Possessing Women*, London: The Women's Press.

²⁰ I am grateful to Leslie Green, Hannah Field, Katherine Simpson, two anonymous referees from this journal, and an audience at the University of Oxford, for comments, criticism, and suggestions.

- Dworkin, Andrea and Catharine A. MacKinnon 1988. *Pornography & Civil Rights: A New Day for Women's Equality*, Minneapolis: Organizing Against Pornography.
- Gates, Henry Louis 1993. Let Them Talk: Why Civil Liberties Pose No Threat to Civil Rights, *New Republic* 37/September 27: 43-49.
- Gilreath, Shannon 2009. "Tell Your Faggot Friend He Owes Me \$500 for My Broken Hand": Thoughts on a Substantive Equality Theory of Free Speech, *Wake Forest Law Review* 44/2: 557-615.
- Goodman, Jeffrey A., Jonathan Schell, Michele G. Alexander, and Scott Eidelman 2008. The Impact of a Derogatory Remark on Prejudice toward a Gay Male Leader, *Journal of Applied Social Psychology* 38/2: 542-55.
- Green, Leslie 1998. Pornographizing, Subordinating, and Silencing, in *Censorship and Silencing: Practices of Cultural Regulation*, ed. Robert C. Post, Los Angeles: Getty Research Institute for the History of Art and the Humanities: 285-311.
- Green, Mitchell S. 2000. Illocutionary Force and Semantic Content, *Linguistics and Philosophy* 23/5: 435-73.
- Greenberg, Jeff and Tom Pyszczynski 1985. The Effect of an Overheard Slur on Evaluations of the Target: How to Spread a Social Disease, *Journal of Experimental Social Psychology* 21/1: 61-72.
- Haslanger, Sally 2004. Oppressions: Racial and Other, in *Racism in Mind*, ed. Michael P. Levine and Tamas Pataki, Ithaca: Cornell University Press: 97-123.
- Heins, Marjorie 1983. Banning Words: A Comment on "Words That Wound", *Harvard Civil Rights-Civil Liberties Law Review* 18/2: 585-92.
- Heyman, Steven J. 2008. *Free Speech and Human Dignity*, New Haven: Yale University Press.
- Hornsby, Jennifer 1994. Illocution and Its Significance, in *Foundations of Speech Act Theory: Philosophical and Linguistic Perspectives*, ed. Savas L. Tsohatzidis, London: Routledge: 187-207.
- Hornsby, Jennifer and Rae Langton 1998. Free Speech and Illocution, *Legal Theory* 4/: 21-37.
- Kelly, Daniel and Erica Roedder 2008. Racial Cognition and the Ethics of Implicit Bias, *Philosophy Compass* 3/3: 522-40.
- Langton, Rae 1993. Speech Acts and Unspeakable Acts, *Philosophy & Public Affairs* 22/4: 293-330.
- Langton, Rae 1998. Subordination, Silence, and Pornography's Authority, in *Censorship and Silencing: Practices of Cultural Regulation*, ed. Robert C. Post, Los Angeles: Getty Research Institute for the History of Art and the Humanities: 261-83.
- Langton, Rae 2012. Beyond Belief: Pragmatics in Hate Speech and Pornography, in *Speech and Harm*, ed. Ishani Maitra and Mary Kate McGowan, Oxford: Oxford University Press: 126-47.
- Langton, Rae and Caroline West 1999. Scorekeeping in a Pornographic Language Game, *Australasian Journal of Philosophy* 77/3: 303-19.
- Lawrence, Charles R. 1993. If He Hollers Let Him Go: Regulating Racist Speech on Campus, in *Words That Wound: Critical Race Theory, Assaultive Speech, and the First Amendment*, Boulder Colorado: Westview: 53-88.

- Leader, Tirza, Brian Mullen, and Diana R. Rice 2009. Complexity and Valence in Ethnophaulisms and Exclusion of Ethnic out-Groups: What Puts the “Hate” into Hate Speech?, *Journal of Personality and Social Psychology* 96/1: 170-82.
- Lewis, David 1979. Scorekeeping in a Language Game, *Journal of Philosophical Logic* 8/1: 339-59.
- Lewis, David 1996. Elusive Knowledge, *Australasian Journal of Philosophy* 74/4: 540-67.
- MacKinnon, Catharine A. 1984. Not a Moral Issue, *Yale Law & Policy Review* 2/2: 321-45.
- Maitra, Ishani 2009. Silencing Speech, *Canadian Journal of Philosophy* 39/2: 309-38.
- Maitra, Ishani and Mary Kate McGowan 2007. The Limits of Free Speech: Pornography and the Question of Coverage, *Legal Theory* 13/: 41-68.
- Matsuda, Mari J. 1993. Public Response to Racist Speech: Considering the Victim’s Story, in *Words That Wound: Critical Race Theory, Assaultive Speech, and the First Amendment*, Boulder Colorado: Westview: 17-51.
- McGowan, Mary Kate 2003. Conversational Exercitives and the Force of Pornography, *Philosophy & Public Affairs* 31/2: 155-89.
- McGowan, Mary Kate 2004. Conversational Exercitives: Something Else We Do with Our Words, *Linguistics and Philosophy* 27/1: 93-111.
- McGowan, Mary Kate 2009. Oppressive Speech, *Australasian Journal of Philosophy* 87/3: 389-407.
- Mullen, Brian and Diana R. Rice 2003. Ethnophaulisms and Exclusion: The Behavioural Consequences of Cognitive Representation of Ethnic Immigrant Groups, *Personality and Social Psychology Bulletin* 29/8: 1056-67.
- Mullen, Brian and Joshua M. Smyth 2004. Immigrant Suicide Rates as a Function of Ethnophaulisms: Hate Speech Predicts Death, *Psychosomatic Medicine* 66/3: 343-48.
- Nielsen, Laura Beth 2004. *License to Harass: Law, Hierarchy, and Offensive Public Speech*, Princeton: Princeton University Press.
- Post, Robert C. 2009. Hate Speech, in *Extreme Speech and Democracy*, ed. Ivan Hare and James Weinstein, Oxford: Oxford University Press: 123-38.
- Richard, Mark 2008. *When Truth Gives Out*, Oxford: Oxford University Press.
- Schwartzman, Lisa H. 2002. Hate Speech, Illocution, and Social Context: A Critique of Judith Butler, *Journal of Social Philosophy* 33/3: 421-41.
- Searle, John. R 1969. *Speech Acts: An Essay in the Philosophy of Language*, Cambridge: Cambridge University Press.
- Steele, Claude M. and Joshua Aronson 1995. Stereotype Threat and the Intellectual Test Performance of African Americans, *Journal of Personality and Social Psychology* 69/5: 797-811.
- Sunstein, Cass R. 1996. On the Expressive Function of the Law, *University of Pennsylvania Law Review* 144/5: 2021-53.
- Taylor, Charles 1985. Language and Human Nature, in *Human Agency and Language: Philosophical Papers 1*, Cambridge: Cambridge University Press: 215-45.

Waldron, Jeremy 2010. Dignity and Defamation: The Visibility of Hate (2009 Oliver Wendell Holmes Lectures),
Harvard Law Review 123/7: 1596-657.