EXCLUSIONARY SPEECH AND CONSTRUCTIONS OF COMMUNITY

A Dissertation
submitted to the Faculty of the
Graduate School of Arts and Sciences
of Georgetown University
in partial fulfillment of the requirements for the
degree of
Doctor of Philosophy
in Philosophy

By

Cassie Monette Herbert, B.A.

Washington, DC
June 6, 2017
EXCLUSIONARY SPEECH AND CONSTRUCTIONS OF COMMUNITY

Cassie Monette Herbert, B.A.

Thesis Advisor: Rebecca Kukla, Ph.D.

ABSTRACT

This project explores the complex ways language, social power, and identity entwine to structure social interactions and the contours of community boundaries. How an audience takes up discursive moves can constitute who is trusted and taken seriously as a knower and who is held in a position of skepticism. How we signal our identity through language moves can impact the epistemic agency an audience takes us to be employing. Derogative terms, such as racial slurs and words like ‘slut’ or ‘lame,’ target social groups, are connected to a history of oppression against that group, and using these terms can structure interactions with group members in complex ways. The words we use can, even unwittingly, subtly alienate and outgroup members of already vulnerable groups. Each chapter draws on speech act theory to serve as an entry point into the extraordinarily complex ways our discursive practices structure our social world and normative interactions.

I begin by examining ‘risky’ speech and how distorted responses to such speech can impact community constitution. To do this, I develop an account of two different kinds of speech acts – accusations and reports – and explore the performative structure of each. In the next chapter I offer a pragmatic account of discursive signaling. Discursive signaling is when we use linguistic moves to call into salience particular aspects of our identity in a given context, which structures how we and those around us understand and navigate those moments and
situations. In the third chapter I develop an expanded conceptual taxonomy of harmful speech. I argue that there is a multidimensional continuum of terms, which I term *derogatives*, that share some or all of the features of paradigmatic slurs. In the fourth chapter, I explore the pragmatics of reclamation projects. I argue that the performative structure of reclamation makes the projects precarious and open to causing harm. In the final chapter, I offer an account of the perlocutionary effects of mentioning slurs. I argue that slurs act as a powerful mechanism for priming pernicious implicit biases about the targeted group. I offer practical guidance for how to navigate talking about slurs.
I am profoundly grateful for things that haven’t happened: I haven’t been told my work “isn’t philosophy,” that feminist philosophy doesn’t matter, or that work about oppression ought to be relegated to the fringes of serious philosophical work. I haven’t felt or been unsupported, and I haven’t had to struggle to be taken seriously as a philosopher. I haven’t received pressure guised as friendly advice to change the focus of my work because of how it might be taken up by the profession or on the job market. And I haven’t felt alienated from faculty in my department because of what I work on. All of this should be the norm – it isn’t something I ought to feel grateful for – but it’s not yet the norm, and it would be very easy to fall into these existing patterns. The reality is that I am deeply grateful for everything I haven’t had to go through. Thank you to everyone for all the daily work it took to avert these too common experiences.

I owe an enormous and heartfelt THANK YOU to my friends and family who have supported and put up with me while I have worked on this. I am especially grateful to all the incredible faculty at Georgetown who have supported, encouraged, and mentored me through this long journey. I sincerely appreciate the hard work and feedback of my committee – Rebecca Kukla, Wayne Davis, Alisa Carse, and Daniel Silvermint: thank you!

Finally, words can’t capture how deeply I appreciate the enormous amounts of emotional and professional support I have received from Alisa Carse, Bryce Huebner, and especially Rebecca Kukla. There really aren’t words. Just thank you.

My deepest thanks,
Cassie Herbert
# Table of Contents

Introduction .............................................................................................................. 1

Chapter 1: Accusations, Reports, And Responding To Risky Speech ...................... 16
  I. Speech acts, reports, and accusations .......................................................... 18
  II. The objective response ................................................................................ 28
  III. Survivors’ speech ...................................................................................... 34
  IV. Objectivity, masking, and legitimizing ....................................................... 44
  V. Conclusion .................................................................................................... 48

Chapter 2: Identity and Discursive Signaling ........................................................... 49
  I. Situated and salient identities ..................................................................... 52
  II. Discursive signaling ................................................................................... 57
  III. Reflexive discursive signaling .................................................................. 61
  IV. Signaling others’ identities ......................................................................... 69
  V. Conclusion .................................................................................................... 73

Chapter 3: Derogatives: Insults, Slurs, and Speech Acts .......................................... 74
  I. Features of slurs .......................................................................................... 77
  II. Derogatives: Expanding the taxonomy of harmful words ......................... 81
  III. Derogatives terms and derogative speech acts ........................................... 96
  IV. Changing words, changing worlds ............................................................ 103
  V. Conclusion .................................................................................................... 106

Chapter 4: The Pragmatics of Reclamation Projects ............................................... 108
  I. Socially embedded slurs .............................................................................. 112
  II. Performative force and discursive distortions ............................................ 117
  III. The performative structure of reclamation ............................................... 122
  IV. The power and limits of reclamation ......................................................... 126

Chapter 5: Talking About Slurs ........................................................................... 130
  I. Puzzles about slurs ...................................................................................... 131
  II. Talking about slurs .................................................................................... 134
  III. Implicit associations .................................................................................. 137
  IV. Slurs as priming words ............................................................................. 141
  V. How to talk about slurs .............................................................................. 147
  VI. Revisiting puzzles about slurs ................................................................... 159

Bibliography .......................................................................................................... 161
INTRODUCTION

We use speech in myriad ways to navigate, construct, and resist the contours of our social world. Speech (verbal, signed, or written) is a medium through which we interact with others – we use discursive tools to convey ideas, construct relationships, and pursue our projects. Discursive interactions permeate our lives, in moments ranging from the dramatic to the mundane. These interactions can have transformative power, in both positive and negative ways. Speech can be welcoming. Speech can be exclusionary. This project focuses on how we use speech to craft, patrol, and renegotiate the boundaries of our communities. We use speech to bring people in to our communities, to welcome, to invite, to build coalitions and express solidarity. Speech can be used to officially induct someone into a community or to subtly cement that they ‘really’ belong through shared jokes and references and playful banter. We also use speech to cast people out of our communities, to condemn, to repudiate, to accuse and to spurn. We use discursive tools such as insults and slurs to label people as having a particular kind of bad character, or as being a certain sort of deplorable person. I explore how we use language to construct ingroups and outgroups, and how this sets up and perpetuates norms for the treatment of particular people. Throughout, I’ll look at how seemingly promising strategies of resistance can unintentionally recreate the harm they were intended to ameliorate.

Our language practices are coupled to matters of social identity. On the one hand, our identities structure what we are able to do with language – a woman may have difficulty receiving uptake as issuing commands, and a Latinx person may not be read as issuing expert speech despite having the relevant training and credentials. At the same time, how we use
language can shape our identities – using the right slang can bring us into or cast us out of communities, and having to beg students to turn in their papers can undermine the speaker’s standing as a professor. Both sides of the connection between language and identity play a role in situating people within or outside of community boundaries.

I am particularly interested in the moments of discursive disconnect, breakdown, and distortion. Much work has already been done looking at the conditions that underwrite successful communication; insofar as failed communication has received attention, it has typically received attention as exactly that: an instance of failure, full stop. J. L. Austin, for example, focused extensively on *infelicities* and *misfires*, the ways in which speech acts might fail to be successful as intended by the speaker (Austin 1962). But while Austin examined the conditions that can lead to these failures, he (and others after him) have paid little attention to what *happens* in these breakdowns. While certainly not all instances of attempted communication break down (or we wouldn’t have communication at all!) I believe it is philosophically fruitful to take as a central point that these disconnects *can* and *do* occur, rather than to relegate them to the fringes of theory. Discursive disconnect and breakdown reveal and create features of our social world, both features related to communication and otherwise; they do not simply elucidate what needs to be in place for successful communication.

Within political and social philosophy, there exists a distinction between ideal and non-ideal theory. Ideal theory is centrally concerned with setting out principles of justice or the good towards which we should strive. It begins from the working assumption that individuals are well intentioned and rational, and that conditions are favorable to achieve
whatever principles are set out. The central project of ideal theory is necessarily abstracted
away from the lived reality of our world; the rationale is that only by first setting out these
abstract ideals can we begin to reason about how to achieve them. Non-ideal theory, on the
other hand, is focused on the lived, embodied, and messy reality of our social world. While it
addresses moving towards a more just society, the project is centrally about attending to and
understanding injustice in our current reality. Non-ideal theory is, to a degree, parasitic on
ideal theory: it builds on a general conception of justice or the good to develop an
understanding of what is going wrong in the status quo. On the other hand, ideal theory runs
the serious risk of being either impracticable because it is so disconnected from our current
social reality, which is rife with injustice and oppression, or being fundamentally flawed
because of the assumptions about human nature and social conditions that it builds from.

I believe there is an interesting parallel to draw here in philosophy of language.
Philosophy of language thus far has largely focused on how things ought to be in order for
successful communication of a speaker’s intent. Some focus has been given to moments of
failed communication, but mostly this has been in the service of elucidating what is
necessary for success; when a speaker is unable to do what they intend with their speech is a
side-issue and is only interesting insofar as it helps us understand successful communication.
That is, the understanding of discursive breakdown is, first, only examined in service of
understanding intentional success, and second, is understood within the framework developed
to understand intentional success. This limits and skews, respectively, our philosophical
understanding of what is happening in these instances.

In contrast to this, I take these moments to failed intentional communication to be rich
and interesting philosophical phenomena in themselves. While I am happy to adopt
successful communication as the aim of discourse, I am centrally interested in the messy reality of how speech acts play out in our social world. This often includes failure. These moments of failure can take a variety of forms, from innocuous mistakes on the part of speakers or audiences, to more structured and predictable

H. P. Grice and Austin developed work that serves as the backbone of traditional pragmatist speech act theory. Feminist philosophers have recently been drawing on their work to address topics such as hate speech and pornography. I’m first going to lay out the basic features of each theory, and then explore how these theories have been updated and applied to issues closely related to my project. Each, I argue, is unsuited to my project here.

Grice argued that we do quite a bit of discursive work via what isn’t said. Grice introduced the notion of implicature: the idea that a speaker conveys information by how something was said rather than merely what was said. He recognized that language is a cooperative project – speakers and audiences work together to produce meaning. He put forward the Cooperative Principle, which states, “Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged” (1989, 26). Audiences, often unwittingly, draw on this principle to make sense of what a speaker is trying to convey. When a speaker includes extra information, the presumption is that this information is relevant even when the speaker doesn’t tag it as such. Similarly, the Cooperative Principle helps us to make sense of seeming failures to answer, such as when a speaker answers the inquiry, “are you going to the party?” with what might literally seem to be a non-sequitur
such as, “I have to work on Saturday.” An audience can easily interpret this as a response to the question, and can fill in the implicature as “I’m not going to the party because I have to work on Saturday. On this view, both what is said and what is implicated are aspects of the speaker’s meaning.

Jennifer Saul explores the ways Grice’s conversational implicatures are normatively structured affairs (Saul 2002). She identifies a distinction between utterer-implicatures, which are those a speaker has tried to implicate, and audience-implicatures, which are those an audience presumes the speaker to have implicated. Though in most cases these align, Saul identifies the moments when they come apart. Saul argues that we can convey unmeant implicatures and that we can say an implicature occurred even when an audience did not properly take up an intended implicature that should have been decipherable. Yet despite her focus on the complexities of implicature and possible disconnects between speaker intent and audience uptake, both aspects here center on the success or failure of communicating the speaker’s intended meaning. I find this approach promising, though unsuited to my project here. I am interested in the normative relations wrought by discursive moves rather than on the success or failure of conveying meaning. For this, I turn to Austin.

According to Austin, the performative force of a speech act is the normative change wrought by that speech act. Austin argues that speech acts have three parts: the locution, or surface grammar of the utterance, the perlocution, or the effects brought about by the utterance, and the illocution, which is the action constituted by the utterance (Austin 1962). Austin recognizes that in many ways, the force of a speech act is outside the control of the speaker; certain felicity conditions must be in place for an utterance to achieve its intended
purpose. To borrow an example from Austin, when a duly designated individual at a ship christening ceremony says, “I name this ship the *Queen Elizabeth,*” that ship is now named the *Queen Elizabeth,* while it wasn’t so before (Austin 1962). However, if “some low type” were to run up to the ship, grab a champagne bottle out of the duly appointed person’s hand, smash it against the ship and say, “I name this ship the *Generalissimo Stalin,*” it would not be the case that the ship is called the *Generalissimo Stalin.* The interloper did not meet the necessary conditions for christening the ship, and the speech act did not have the intended pragmatic force of christening the ship with a new name. Austin terms this a *misfire.* Misfires help us to understand that conditions that need to be in place for a speech act to be successful. However, he does not further investigate these moments of disconnect between what a speaker intended to do and what they did, in fact, bring about. Rather, a misfire is deemed a moment of failure and that is the end of the story.

Rae Langton builds on Austin’s speech act theory to develop an account of misfires that occur due to how social norms impede on the intended uptake of speech acts that should, ostensibly, be successful. Langton terms this *illocutionary disablement* (Langton 1993). Here, a speaker is unable to receive uptake as performing the intended type of speech act due to alternate norms that differ from the publicly recognized norms. Langton considers pornography in order to illustrate this phenomenon: pornography\(^1\), she claims, constitutes social norms by which women’s sexual refusals cannot be properly recognized as such. On this account, pornography helps to establish and perpetuate social norms under which

\(^1\) Langton has in mind a particular kind of pornography which might best be described as “mainstream porn.” This kind of pornography is straight, body normative, oriented towards male viewers, and depictures subordinating representations of women.
women lack sexual agency. These norms set discursive conventions whereby women’s sexual refusals are unintelligible as refusals. It’s not simply that the refusals are unheard or ignored, but rather that they aren’t recognized as refusals at all. When this occurs, Langton argues, women have been *silenced*.

Langton highlights something important when she argues that in these cases a speaker is unable to do what she intends with her words. However, by relying so heavily on the Austinean framework she is blurring together two importantly distinct kinds of illocutionary disablement: illocutionary *silencing* and illocutionary *distortion*. For the most part Langton, and the many others who have taken up her work, implicitly focuses on illocutionary *silencing*. In these cases the speech act isn’t taken as any kind of illocutionary act at all. It is simply a misfire; a failed act. In the case of sexual refusals, the speech act is neither unheard nor ignored, yet it doesn’t receive uptake from a hearer as anything – it’s akin to screaming into the wind.

However, speech can still do things – often powerful and important things – even when the changes wrought by the act are not those intended by the speaker. These are instances of what I term illocutionary *distortion*: when the audience takes up a speech act as a different type of illocutionary act than that intended by the speaker. To return to the case of sexual refusals, pornography can set up gendered norms in which women are understood to ‘play hard to get’ as a form of sexual encouragement. In these contexts, a woman’s sexual refusal isn’t taken as *nothing*, but instead as a kind of encouragement and participation. That is, the force of her speech is coopted and turned against her, and to claim she is *silenced* paints an incomplete picture; it ignores the real and harmful pragmatic force of her words.
To have her speech turned against her is an importantly different kind of disempowerment than to not be recognized as speaking at all. To term a sexual refusal issued in a context where women aren’t recognized as having sexual agency as a misfire is grossly inaccurate and incomplete. Yet within the Austinean framework this is the only conceptual tool available. The Austinean framework doesn’t account for the continued force of speech acts when the intent does not align with the uptake the act receives. For this reason, Langton’s account is also fundamentally limited. We need recognition of both illocutionary silencing and illocutionary distortion in order to have an appropriately rich conceptual framework of the ways in which illocution can fail.

Most of the time it’s clear how a speech act ought properly be taken up and the speaker’s intent aligns with the uptake it receives. Sometimes, however, a hearer must navigate competing discursive norms: As Rebecca Kukla argues, different sets of norms yield different, yet normatively appropriate, understandings of the illocution performed (Kukla 2012). So, to return once again to Langton’s example, explicit norms about the right of sexual refusal yield one kind of uptake of a woman’s intended refusal, while the social norms constituted by subordinative pornography yield an entirely different result. Sometimes features of a speaker’s identity will do this work of calling to salience one set of discursive norms over another (Kukla 2012). Here, the woman’s identity as a woman intersects with the norms constituted by the pornography to lead a hearer to take up her refusal as encouragement. The point is that both official and unofficial social norms and practices work to govern our discursive conventions, and these various sets of practices can come into conflict. Here, it substantially matters which norms the hearer is being guided by. In these
cases where the governing discursive conventions are underdetermined, distortions are especially likely.

Starting from a position in which breakdown is not only possible but important allows us to better see the various ways that speech acts can fail to achieve the desired end. These failures are not homogenous; they can be varied and complex in themselves, and these unique failures lead to all sorts of significant outcomes. In the case of illocutionary disablement, both silencing and distortion are philosophically rich and interesting, and deserving of our attention. Moreover, attending to discursive distortions helps to illuminate the complex interplay between competing sets of social norms.

Furthermore, this non-ideal approach to philosophy of language falls within a larger trend in philosophy to attend to epistemic and discursive injustice. The traditional ideal theory approach to philosophy of language privileges the intent of the speaker in both its focus and conceptual understanding. Historically, it’s been those who have social power — especially white, cis, straight, able bodied, affluent men -- who have been recognized as speakers. The non-ideal approach I favor doesn’t privilege either the speaker or the audience, and this can be understood as a move away from privileging only the discursive agency of members of the group that we have historically been attending to. Rather more interestingly, this approach recognizes the multiple, complex, and co-existing discursive and social norms at play, and doesn’t give priority to only the set of norms that is recognized and drawn on by the group with social power. That is, it centers what José Medina refers to as “polyphonic contextualism” — multiple, simultaneous discursive communities (Medina 2006, 2012). By taking as a primary focal point the moments of breakdown and disconnect, this approach essentially recognizes not only differing social positions, but complex sets of social
and discursive norms that supervene on, cross between, and overlap these positions --- while not prioritizing one over the other.

I make use of the approach to philosophy of language put forth by Rebecca Kukla and Mark Lance in *Yo! and Lo! The Pragmatic Topography of the Space of Reasons*. Essential to their view is that, “language has systematic normative effects and functions, and that these essentially depend on the concrete ways in which speakers are enmeshed in social communities and environments, and that discursive performances systematically transform the normative statuses of speakers and of those they are spoken to” (Kukla and Lance 2012, 2). This approach focuses on the material effects of our discursive interactions, and the transformative power these interactions can have. Their approach is particularly well suited for my project; though their book doesn’t focus on the non-ideal moments of philosophy of language, it lends itself well to this project by focusing on the normative effects of language rather than on speaker intent and meaning, and by recognizing the multiple, complex, and perhaps overlapping social contexts in which speech is deployed.

Kukla and Lance focus on the inputs and outputs of speech acts – the normative conditions necessary to issue a given type of speech act and the normative effects it has on an audience. They explore the performative structure of various *categories* of speech acts – what differentiates, say, a declarative from a promise. Most of the time, an individual speech act clearly fits into a given category (even if it inhabits more than one category at once): the speaker meets the necessary input conditions, they intend to issue that type of speech act, and the audience takes it up as such. However, I am especially interested in moments of disconnect and breakdown, moments where intent and uptake do not align. Here it’s helpful to pull apart the performative structure of types of speech acts, the speaker’s intentions in
issuing this particular speech act, and the audience’s uptake of that speech act. It can be ambiguous which category an utterance fits into, especially when it meets the input conditions for more than one type of speech act, the speaker intends it to be one thing, and the audience takes it up as another. While the audience’s uptake determines the material effects of the speech act, it’s philosophically fruitful to attend to disconnects between intent and uptake. I attend to how the performative structure of different categories of speech acts can shed help illuminate what is going on in these moments of distortion.

The first chapter, “Accusations, Reports, and Responding to Risky Speech,” looks at ‘risky’ speech and how distorted responses to such speech can materially impact community constitution. Risky speech is speech that challenges entrenched systems of power in some way and which carries with it real risks (epistemic, social, or physical) to the speaker. I focus in particular on accounts of racism and of sexual violence as paradigmatic examples of risky speech. In order to make sense of how this speech is commonly taken up, I develop an account of two different kinds of speech acts – accusations and reports. Though on the surface these speech acts seem quite similar, they have distinctly different performative structures. I will lay out each of them in turn. I argue that risky speech is commonly taken up as an accusation – whether that is what it was intended by the speaker or not. This does not merely distort the speaker’s intentions; it warps the responses to their speech in ways that have significant material consequences. This kind of distorted uptake facilitates resistance to risky speech by serving as a masking and legitimizing move. It is a way to dismiss the risky speech while still allowing the audience the veneer of responding properly and attentively.
In the second chapter, “Identity and Discursive Signaling,” I explore how we use language moves to make salient aspects of identity. I offer a pragmatic account of what I term *discursive signaling*. Discursive signaling is when we use linguistic moves to call into salience particular aspects of our identity in a given context, which structures how we and those around us understand and navigate those moments and situations. I begin by discussing the complexities of social identity. Then, I lay out the pragmatics of discursive signaling as a kind of speech act. Slurs, I argue, are particularly powerful tools for discursive signaling. I explore two different ways we do this with our own identity: signaling directed at other ingroup members, as a way to constitute a collective “we,” and signaling directed at outgroup members, to call on outsiders to attend to an aspect of our identity they may have otherwise failed to regard as significant in that moment. In the fourth section I explore the ways we call aspects of others’ identities into salience, and the difference this makes in how we understand and interpret their actions, and how this kind of signaling can play a role in constructing their social identity.

In the third chapter, “Derogatives: Insults, Slurs, and Speech Acts,” I expand the conceptual taxonomy of harmful words. First, I lay out the generally agreed upon features of slurs. Second, I offer expanded conceptual machinery for thinking about slurs. I argue that there is a multidimensional continuum of terms, which I term derogatives, that share some or all of the features of paradigmatic slurs. Just how a derogative is located on this multidimensional continuum impacts how it functions in the world. Third, I lay out the connections between derogatives and speech acts. Derogatives are paradigmatically deployed in derogatory speech acts. But not all derogatory speech acts involve derogative terms, and derogative terms are deployed in many speech acts other than derogatory ones. Finally,
based on this expanded conception of harmful speech, I focus on the various ways derogatives can enter into and go out of our lexicon: via linguistic shift in which derogatives become obsolete, the “euphemism treadmill” in which new words come to be slurs because of stigma about the associated group, or through intentional reclamation projects.

In the fourth chapter, “The Pragmatics of Reclamation Projects,” I present the performative structure of reclamation projects; because of the overlapping discursive communities in which participants are situated, these projects are particularly precarious and risk reinforcing the very norms they are intended to subvert. Derogatory terms can be powerful mechanisms of subordination, while re-appropriating these terms can be a strategy to fight back against social injustice. I argue that projects seeking to reclaim slurs have a performative structure that raises particular hazards. Whereas more familiar forms of protest may fail to bring about their intended result, attempts to re-appropriate slurs can fail to be understood as transgressive acts at all. When attempts at reclamation fail, their force is distorted; context and convention lead the hearer to give uptake to the speech act as a traditional deployment of the slur. The force of this traditional use is to validate and re-entrench the very norms the act was intended to subvert. This is the precarious structure of reclamation projects: when successful, reclamation is the subversion of powerful mechanisms of oppression, but when unsuccessful, the act has the ironic force of constituting mechanisms of oppression.

In the final chapter, “Talking About Slurs,” I continue to explore unintended effects of speech. Here, I offer an account of the perlocutionary effects of mentioning slurs. While it is generally well recognized that using slurs is wrong, many are still unsure just how to navigate talking about slurs. I argue that even in cases when slurs are merely being mentioned, though
not used, this is harmful. I draw on empirical studies to argue that slurs act as a powerful mechanism for priming pernicious implicit biases about the targeted group. Moreover, there’s reason to believe this holds even in cases where the slur is merely being mentioned. Based on this, I offer practical guidance for how to navigate talking about slurs. Finally, lay the foundations for how this priming effect of slurs sheds light on the classical puzzles about slurs.

Together, this project explores the complex ways language, social power, and identity entwine to structure social interactions and the contours of community boundaries. How an audience takes up discursive moves can constitute who is trusted and taken seriously as a knower and who is held in a position of skepticism. How we signal our identity through language moves can impact the epistemic agency an audience takes us to be employing. The words we use can, even unwittingly, subtly alienate and outgroup members of already vulnerable groups. Derogative terms, such as racial slurs and words like ‘slut’ or ‘lame,’ target social groups, are connected to a history of oppression against that group, and using these terms can structure normative interactions with group members in complex, multivariable ways. And efforts to dismantle derogative terms via reclamation can be both empowering and hazardous, to different degrees, depending on the social locations of those involved. Each chapter serves as an entry point to the extraordinarily complex ways our discursive practices structure our social world on the basis of identity and community, and vice versa.

By centering moments of breakdown and disconnect throughout, this project approaches the constructions of identity and community from a recognition of the multiplicity of social norms at play, and the messy ways these can interact. The project takes seriously the unintended harm that can occur from seemingly positive or neutral discursive
acts – from the way taking accounts of wrongdoing as accusations can “mask” moves to undermine the speaker, to the surprising harm from merely mentioning slurs, to the unintended reinforcement of pernicious norms that can occur in reclamation projects. Whether intended or not, these harmful effects can serve to perpetuate exclusion and marginalization of the people they target, which is particularly dangerous when directed at already vulnerable groups of people.

In this project I seek to recognize the discursive power of a multiplicity of speakers and audiences. The non-ideal approach to philosophy of language provides space to attend to the agency of those who are historically not recognized as discursive participants and to the norms that it hasn’t been in the interest of those with social power to recognize explicitly. These are especially essential methodological moves in a project about who is constituted as a community insider. By giving conceptual priority to the complex, overlapping, and sometimes conflicting discursive norms and epistemic resources of a range of people and communities, I hope to offer a more robust account of how the interplay of language and identity structure our daily lives.
It’s appalling the way her case has been handled thus far. Still it’s important to note that the young man in question is innocent until proven guilty. I want to see her case go to trial and be judged on its own merits, but we must not jump to conclusions about the player’s guilt… I’m just trying to put some common sense into people who drag others name through the mud online before all the facts have been heard. Now the fact that she did a rape kit and is willing to prosecute speaks very strongly in her favor. But the people calling him a rapist, before he has been convicted, are just as bad as the people who call her a liar based on her alcohol-level and clotting. We should always help the victim, but we as a society cannot judge nor punish unless a person is convicted beyond a reasonable doubt.

- Comment on a woman’s op-ed about being raped by a fellow student (Marans 2016)

The above comment illustrates an increasingly common form of response to accounts of injustice. I term this the ‘objective’ response. The objective response calls for objectivity and impartiality, for caution, for evidence to be presented and examined, and for waiting to pass judgment until both sides have been heard. Though this response mirrors legal responses to accounts of wrongdoing, here I am focusing on how this response is deployed in daily life outside of legal contexts. This response occurs in public discourse, in institutional settings, in online comments, and in private interactions, often from well-intentioned people. On the surface, the objective response seems like an appropriate move – it is a way to take the speaker seriously and the call for objectivity appears as a measured and even wise reaction. Yet, as I argue, as a default response it has significant problems.

The objective response takes the speaker to be performing a particular action – accusing – though often this is distinctly not what they intended to do when they came forward. As
such, it distorts the speaker’s intentions, and blocks their words from being taken as the report (or other act) the speaker intended it to be. How their speech is taken up structures both the pragmatics of what they are taken to be doing, as well as the material steps taken as a result of their account. Though the objective response appears impartial, in practice it prioritizes the account of the identified wrongdoer rather than the speaker who is disclosing their experiences. Moreover, this response can function to obscure more well-known forms of dismissal, such as when an audience doesn’t take a speaker to be credible on the basis of their race, gender, or other social identity. Even when the audience has good intentions, the objective response can be a masking move, giving the appearance of an apt and impartial response while pragmatically serving to subtly dismiss the speaker’s account.

Sally Haslanger has written on how objectifying reason has the veneer of impartiality and aperspectivity, yet masks the ways in which social facts are produced and reinforced by this epistemic engagement (Haslanger 2012). Similarly, the objective response uses its veneer of impartiality to obscure how the response itself serves to constitute the normative standing of the speaker by casting them as ‘appropriately’ trustworthy or not, as a member of the moral community or someone relegated to the margins. The objective response does all of this with the veneer of taking the person’s speech seriously and responding appropriately.

I utilize speech act theory to elucidate the pragmatics of the objective response. In order to do this, I develop an account of two different kinds of speech acts – accusations and reports. Though on the surface these speech acts seem quite similar, they have distinctly different performative structures. I will lay out each of them in turn. I then argue that the objective response takes up accounts of wrongdoing as accusations – whether that is what they were intended to be or not. Though it is an apt response to an accusation, it is not an
appropriate response to a report. This distorted uptake is especially likely to occur in response to *risky speech* – speech that challenges the status quo. This doesn’t merely distort the speaker’s intentions; it warps the responses to their speech in ways that have significant material consequences. To show this, I explore the role of the objective response in uptake of survivors’ accounts of sexual violence. Finally, I lay out how the objective response can facilitate resistance to risky speech by serving as a *masking and legitimizing* move. It is a way to dismiss the risky speech while still allowing the audience the veneer of responding properly and attentively.

I. Speech acts, reports, and accusations

Speech *does things*. As J.L. Austin and others have pointed out, we do more with speech than simply put forward information. Speech is an action, and through this action we make real changes to the world. We can command, invite, entreat, persuade, call to order, or any number of things via speech. Each of these actions has a particular performative structure. To start with a familiar example, a speaker must have the appropriate standing to issue a command and be issuing it to the right people, in the right context, for it to succeed. A student ordering a professor to return their exam will simply fail, while that professor can command the student to turn in their assigned paper. But the professor cannot command a student who is not in the course to turn in a paper and they cannot arbitrarily order their student to wear purple, eat bananas for breakfast, or any number of things the professor simply has no authority over. While commands require certain authority conditions to be felicitous, as Rebecca Kukla and Mark Lance have pointed out, other speech acts are agent-
neutral in their inputs – anyone can properly issue them, regardless of whether or not they have authority in that context (Kukla and Lance 2009). Anyone can issue declaratives, for example – it doesn’t take a special status to do so.

Not only are speech acts distinct from one another based on their entry conditions and to whom they can be directed, but they differ by their function and what they do. To return again to commands, a command, issued with the proper authority, in the right context, and to the right person, creates a requirement for that person to do something. A boss commanding their employee to attend a meeting properly expects the employee to comply; if they do not, they have responded inappropriately. (Of course, sometimes life happens and we miss meetings – but when that happens we offer excuses and apologies in recognition of having failed to do something we ought to have done.) On the other hand, an invitation creates an option which one can take up or not, and both are appropriate responses. An invitation to coffee between friends does not automatically require attendance – the purpose of an invitation is that it is optional, and we would be dismayed to find out a friend had come because they felt they had to. Yet an invitation is not a neutral laying out of the options either: someone has offered us something by extending the invitation and they are properly owed gratitude, however minimal, for doing so (“Thanks for the invite but I just grabbed coffee a bit ago!”) (Derrida 1966/1978). An invitation is not simply a weak command or one issued by someone with only marginal authority. Invitations and commands have different performative structures and pragmatic outputs, even when they both aim at the same result of getting someone to show up for a particular event. Even when the effects are similar, the particular speech act performed makes a difference in why the
person is there, whether they ought to thank the other for the invitation, the tenor of their ongoing interactions, etc.

Here, I want to explore two speech acts that have received little direct philosophical attention: reports and accusations. Accusations seek to hold a person accountable for some sort of wrongdoing, while reports are declaratives about something the speaker has experienced, observed, or investigated. Each has a distinct performative structure even though their surface grammar can at times be interchangeable.

Reports help make up the fabric of our daily lives. “I woke up and made coffee,” is a familiar report about the start of the speaker’s day. This is a type of declarative, a claim about publically accessible truth (Kukla and Lance 2009). Reports are truth claims about something the speaker has particular standing to talk about, i.e., the things the speaker experienced or investigated. Presumably, anyone in a similar position would assent to the report, and so the truth claims offered are publicly accessible. The difference between reports and declaratives lies in the relation between the speaker and the subject. Reports require a first-personal relation between the speaker and the topic on which they are speaking. If a person reads a book about sailing, they may give a report on that book, but they cannot report on sailing itself – they must either go sailing or study up and become an expert on sailing to be able to report back on it. In contrast, that speaker is warranted in issuing declaratives about sailing on the basis of the book alone (though of course their statement might be wrong.)

Standing to issue a report may come from officially recognized experience – a botanist may give a report about the growth rates of the trees she has been studying, or a journalist may report on police corruption they have been investigating. Often though this standing is more informal and gained in virtue of the content being about the speaker’s life and
experiences. A person who spends much of their free time playing video games is warranted to issue (some) reports about gaming culture, even though they do not have any professional credentials in that area, and a Black Lives Matter activist may well have the standing to issue (some) reports about implicit racism in virtue of what they have experienced in their daily life and the contextual knowledge they gained as a member of this activist community.

Reports constitute normative changes in their audience, and they do this in two interconnected ways: they are an invitation to trust the speaker, and on the basis of this they call for belief from the audience. Reports, like tellings (declaratives told to a particular person or persons), are an invitation to trust the speaker and they call on the audience to take up this invitation and the attendant normative relations (Kukla and Lance 2009; Moran 2005). That is, reports instantiate a particular second-personal relationship between the speaker and the audience (Kukla and Lance 2009: 167). This holds even in such formal contexts as issuing the findings from a scientific investigation: the speaker here is calling for the audience to trust her, both in what she says and that she really did investigate the subject as she ought to have. Reports call for belief on the basis of trust, and proper uptake of a report includes believing what the speaker has said.

As Richard Moran has argued, in telling, a speaker offers the one told their word, and vouchsafes the truth of the content of the telling (Moran 2005). Reports take this a step further, assuring the audience that the speaker in some way personally experienced or looked into the issue at hand. That is, reports are an invitation to trust, both in terms of trusting that the content of the account is true and in terms of trusting in the speaker themself in terms of their first-personal witnessing of that content. This raises the stakes of a report, and disbelief in the report is subsequently a stronger repudiation than it is for tellings generally. A person
may fail as a knower by, say, being too gullible and telling others unwarranted and false information, but failure as a reporter means that person failed to aptly perceive their own lived experiences.

The proper uptake of a report is to believe what was said. The audience may notice or look for conflicting evidence, and if any is found they ought to reevaluate their belief, but the point is that it is improper to automatically presume a report is suspect (barring a history of lying or misperceptions from the speaker). While an audience may well ask for more information in response to a report, it is not because more evidence is needed in order to believe the account – it is usually just to find out more about what happened. When a spouse announces, “I had an awful day,” and receives “Oh no! What happened??” in response, it is not because that spouse was not believed about the awfulness of their day – it is because their partner wants to know more in order to properly commiserate.

Of course, this does not mean that all reports are necessarily true – sometimes the speaker is lying, misinformed, or there may be legitimate differences of opinion in interpreting information. Sometimes we find conflicting evidence – that the trees really aren’t growing as reported, that the gamer only plays in a highly esoteric community that does not generalize to the rest of gaming culture, or that our partner has been lying about their day – and these can and should cause the audience to re-evaluate their belief in the report. But it is normatively improper to start from this position of distrust. Rather, the performative structure of reports is to enter into this second-personal relation of trust and on the basis of this grant belief in the account offered, and if conflicting evidence emerges to then re-evaluate that belief. Until and unless there is such conflicting evidence, the report itself gives
reason to believe the account of the speaker and to incorporate their account into our decision making as we go forward.

Accusations, on the other hand, aim to hold someone to account for their wrongdoing. The wrongdoing can range from the trivial, ‘You took my pen!’ to the serious, “She sexually assaulted me.” While accusations can be leveled by the person wronged or by another, and can be directed to the accused directly or offered to a third party, at their core they seek to hold the accused accountable for what they have done.

Accusations are about wrongdoing. This is definitional: a statement simply isn’t an accusation unless it is about norm violation in some way. “You said hello when I greeted you!” just does not work as an accusation (barring some atypical contextual norm about proper greetings). Proper uptake of an accusation calls for the accused to feel the pull of the claimed norm violation in some way. A person who simply says, “Yup. I did that,” is either not properly taking up the speech act as an accusation or they are rejecting the grip of that norm on their behavior. However, there can be and are all kinds of contextual norms whose breaking can warrant an accusation: “You were supposed to rescue me when he started talking about all his cats!” may track a particular intersubjective norm between friends that does not hold more broadly in others’ social interactions. Or, what may appear on the surface to be an accusation may simply be a declarative if the norm in question is not followed by the conversational participants: “You stole that!” uttered amongst thieves is unlikely to be an accusation, but instead a congratulation of a job well done. If the speaker said it accusatorily, the other would likely be simply confused – after all, stealing is what they do. So while just which norms are the focus of the accusation may vary depending on
context, accusations are a way to call attention to those norms and the accused’s failure to properly follow them.

However, accusations do more than point out wrongdoing; they are designed to hold the accused accountable for their behavior. Accusations are a way to initiate the practice of holding others responsible (though they may not be the only entry point into this practice). Accusations seek to secure apology, recompense, recognition, or revenge for whatever wrongdoing has occurred. As Coleen Macnamara points out, “The internal aim of these sanctioning behaviors [punishments or reproof] is to induce what we might call first-personal practical uptake of the ought-violation in the one we’re holding accountable—to get the wrongdoer to acknowledge her wrongdoing, feel remorse, apologize, make amends, and commit to doing right in the future. It achieves this aim by imposing burdens—the pain of punishment, the sting of reproof” (Macnamara 2011, 90).

In this sense, accusations are backward-looking; they seek to right a state of affairs in some way. While accusations about relatively trivial things may simply require an apology (“Oh no, I’m so sorry I forgot to do my chores!”), more serious violations may require community intervention to properly hold the person to account—when the stakes are high, the speaker may not be positioned to do it themself. Others are brought in to help—perhaps to fire the accused from the workplace whose rules they violated, or by publically taking the accused to task for what they did and seeking recognition that it was wrong. So accusations can be structured either second- or third-personally. Second-personal accusations, (“You didn’t take out the trash! You always leave the housework to me”) are directed to the person who is claimed to have done something wrong. Third-personal accusations, (“She did it. She’s the one who mugged me”) are directed to others about the alleged wrongdoer.
Fred J. Kauffeld proposes that accusations are paradigmatically second-personal, with third-personal accusations being “more complex variants on the basic paradigm” on the normal model (Kauffeld 1998, 253). Kauffeld argues, “It can be true that an accusation was made even though no third party judge was addressed, but it cannot be true that an accusation has been made in the absence of an overt intention to secure an answering response from the accused” (Kauffeld 1998, fn 16). This seems wrong to me. Accusations are often leveled against those the speaker knows will never seriously engage with the accusation, perhaps because they will not admit of wrongdoing, or do not feel the grip of the norm they violated. While the accused may issue a denial, this is, at best, a surface-level response – the denial here is a way for the accused to brush off the accusation. In these cases, the speaker enters into the accusation knowing they will not receive a real response from the person they are accusing. Yet they do so anyway, by issuing the accusation to others and with the intent that they, collectively, can hold the accused to account even if that person will never deign to answer. Sometimes the whole point is for there to be a public record of the accusation, supporting evidence, and eventual ruling; and the only outcome is a public acknowledgement that the accused was guilty of the wrongdoing. For example, during the 2016 presidential election, at least eleven women came forward with claims that Donald Trump had sexually harassed them (Blau 2016). These women had little reason to think that Trump would admit to wrongdoing or even, given his statements regarding women’s bodies, feel the grip of the norms he was accused of violating. These third-personal accusations weren’t directed at securing a response from Trump, but rather as a call to the public to hold him to account.
The proper response to an accusation is to hold off on belief unless and until supporting evidence has been produced. This is because accusations are a way to hold someone to account, which is a particular kind of normative action. To hold someone responsible is a deontic move; it calls for the person to recognize that they have not properly taken up the *oughts* that bind them (Macnamara 2011). Consider: not taking out the trash is, itself, a relatively trivial thing. Not taking out the trash when you ought to have, though, is a violation of the commitments you have made to your housemate. It is a deontic failure, not simply a procedural one. An accusation calls for the audience to be ready to engage in this normatively weighty action of holding someone responsible (if issued third-personally) or of taking on responsibility (if issued second-personally). This kind of holding to account ought not be done lightly or cavalierly. In virtue of this, accusations require *support*, paradigmatically in the form of corroborating evidence. This is part of what distinguishes accusations from other sorts of speech acts. Unlike declaratives, reports, or promises, which the audience can believe in virtue of the speech act itself, accusations ought not succeed in eliciting belief on their own. The content of an accusation ought not be automatically accepted as true; its content is *provisional* until it is supported, and then the accusation is retrospectively taken as well-founded or not. As such, part of the performative structure of an accusation is for the audience to place the testimony in a suspended status of belief, and to hold off on believing that the accused has in fact committed what was claimed. Corroborating evidence is sought, and if any is found then this lack of belief is reevaluated. Part of responding appropriately to an accusation is to hold off on automatically presuming its content to be true; this is a crucial part of their performative structure. Those who automatically believe what was said in the accusation are jumping to conclusions and not adhering to the proper uptake of an
accusation. Unless and until corroborating evidence is proffered, the audience does not (and ought not) incorporate the propositional content of the accusation into their beliefs.

Accusations require evidence in order to be believed. More than this, they require corroborating evidence, and so they invite the audience to scrutinize whatever supporting evidence is proffered. A speaker is typically granted time to find such evidence and it may be offered as it is discovered, or the speaker may have it in hand as they issue the accusation. This invitation to scrutinize the evidence helps differentiate accusations from, say, ultimatums. A spouse brandishing illicit text messages when accusing their partner of cheating is holding them to account for violating the norms of their relationship, while at the same time leaving (at least potential) space for the partner to dispute or contextualize the text messages. The texts are being used as supporting evidence, and so whether they in fact corroborate the accusation is specifically open to discussion. In contrast, the spouse who brandishes the text messages and declares “It’s over,” isn’t necessarily making any move to hold their partner responsible for breaking their commitments, and though the text messages may be the reason for the ultimatum, the spouse isn’t presenting them as evidence and isn’t inviting scrutiny of their relevance.

Accusations and reports differ, both in terms of what sort of normative relation is initiated and instantiated by the speech act, how the audience is invited to respond to evidence, and whether the audience ought to believe or remain agnostic about the content of what was said. Yet these speech acts are often indecipherable at the level of surface grammar. “She raped me,” could fall into either category. This helps make it easy for intended reports to be taken up and treated as accusations, and vice versa.
II. The objective response

Now, we are in a position to better understand what is going on in the objective response. This response takes up the speaker’s account as an accusation. This is illustrated throughout the response. This response calls for caution, which is appropriate when the normative stakes are high, as they are when holding someone to account. The objective response calls for evidence to be produced to support the speaker’s account, which is needed at the outset for accusations though not for reports. More than that, it explicitly urges holding off on belief in what was said until and unless such evidence has been produced and examined. This would be an inappropriate response to a report, though an apt one to an accusation. The objective response is a clear way to demonstrate uptake of the speech as an accusation. But more than that, the objective response is the paradigmatic appropriate response to an accusation. A person who gives this response to a (third-personal) accusation has responded properly, correctly, and ought to be lauded.

This is all well and good when the objective response is given to a speech act that was issued as an accusation. However, this response is also given to speakers intending to issue reports. As such, this warps what the speaker is able to do with their words and the normative moves they are able to bring about through those words. This distortion isn’t merely random or bad luck. Nor is it a one-off occurrence. Risky speech – speech that pushes against the status quo – is particularly well-situated to receive this response. The objective response is a typical or default response to this kind of speech. This blocks the speaker from being able to do with their words as they ought to be able to do.
Speech acts can be identical in their grammatical construction even when they are distinctly different actions. Accusations are often indecipherable from reports on their surface grammar alone. “She took my pen!” may be leveled as an accusation intended to elicit an apology and the return of the pen, or it may be part of the speaker reporting on their day to a friend. We rely on complex sets of discursive conventions to govern how we take up such speech acts. Generally, audiences succeed in correctly giving uptake to the act as it was intended, but as Rae Langton, Rebecca Kukla, and others have pointed out, sometimes alternate sets of norms are made salient by a particular context or aspects of the speaker’s social identity and this can lead to frustrating, harmful, or dire results. A well-known example of this is when “no”s are eroticized and ‘playing hard to get’ is expected of women in a sexual encounter, which can lead an audience to give uptake to a woman’s “no” as a move within that ongoing sexual interaction rather than as a denial or revocation of consent (Herbert 2015; Kukla 2012; Langton 1993). When these conventions systematically lead an audience to take up the speech act in such a way that furthers the speaker’s disempowerment on the basis of their social position, this is an act of what Rebecca Kukla terms discursive injustice (Kukla 2012).

While not all reports are about wrongdoing, at least some of them are. And this makes accusations and reports particularly difficult to sort out. It is only through careful attention to context, to who is involved in the conversation, and to other subtle discursive clues that these utterances can receive the intended uptake. When reports are taken up as accusations, or vice versa, this distorts the pragmatic force of the speech; the speaker is not able to do with their speech as they wish and ought to be able to do. When reports and accusations are mistaken for one another, this limits what the speaker can do with their speech in that
moment, and distorts the normative relation the speaker is able to bring about with their speech. Someone saying “You didn’t take out the trash again” may merely be intending to helpfully remind their roommate of their unmet chores, only to find themself in the unwanted position of holding them accountable for these lapses. When this happens as a one-off mistake, it is frustrating and can lead to a variety of interpersonal tensions. However, often such misconstruals are not merely random misunderstandings, but instead part of a pattern of systemic injustice. How we take up speech acts is structured by entrenched norms that help to maintain social hierarchies and systems of power.

Speaker identity and context can play important roles in how accounts of wrongdoing are taken up. When a speaker is socially privileged, this privilege can skew how both their reports and accusations are taken up. A privileged speaker may find their reports more robustly trusted – and by extension, their experiential agency more valued. And a privileged speaker may find their identity playing a significant role in how much, if any, corroborating evidence is needed for their accusations to be taken as well founded. A wealthy white man can more easily initiate the process of holding someone to account – say, by recruiting legal resource or social sanctions – in part because his word can stand as sufficient evidence on its own.

Social privilege is a complex and contextual phenomenon. For example, white women have historically been marginalized and deemed not to have credibility, be it because they’re too emotional, hysterical, incapable of reason, or simply ignorant. Reports of sexual violence from these women have often been met with indifference or skepticism. An exception, however, is when a white woman accuses a black man of rape; audiences believe these accounts much more readily than other accounts of sexual violence. US history is rife with
examples. When Carolyn Bryant, a white woman, accused the black 14-year-old Emmett Till of sexual harassment, not only was she believed but her speech prompted the horrific torture and murder of a child. Bryant, decades later, admitted the harassment never happened (Pérez-Peña 2017). But because she inhabited a position of privilege relative to a black adolescent, and because she accused him of violating a deeply entrenched social norm that protected white women from racial outsiders, her speech wasn’t met with mistrust or with default skepticism. The white men of her family were far too ready to believe that such harassment occurred and to take her at her word.

While privilege can bolster a speaker’s trustworthiness, marginalization can do the opposite. When a speaker is a member of a marginalized group, they may be by default held in a position of mistrust. When a speaker from a marginalized group issues a report, their invitation to trust may go unanswered. Taking their speech as an accusation can be an insidious way to sidestep the invitation to trust that is part of a report. And then they may be required to produce (if they can) additional evidence in order for their accusation to be considered well founded. In addition to the social power of the speaker, the content of the speech can impact how it is taken up.

Risky speech is especially likely to be taken as an accusation rather than as a report. Risky speech is speech that challenges dominant norms and practices, and in virtue of this can make the speaker vulnerable to certain jeopardizations. These jeopardizations are systemic.

---

2 ‘Though I am borrowing Dotson’s term here, I am broadening her account of ‘risky speech.’ According to Dotson, risky speech is speech whose content the audience “may fail to find fully intelligible” (2011: 244). I don’t disagree that this epistemic risk is one hallmark of risky speech, but also want to acknowledge that this kind of speech carries with it manifest social, professional, and physical risks to the speaker. And indeed, Dotson later identifies these harms as part of the risks of risky speech, so we may agree entirely.'
and stable. They may take the form of epistemic or discursive risk, in which the audience
fails to comprehend the content of the speech or takes up the speech in such a way that
further disempowers the speaker. In a famous example, Kimberlé Crenshaw points out how
black women who speak out about domestic violence within their communities are taken to
be reinforcing existing pernicious stereotypes about “violent” black men, rather than as
seeking help and resources for a problem that spans communities (Crenshaw 1997). Risky
speech also carries with it social and professional hazards. These hazards might include
friends, family, or employers distancing themselves or turning on the speaker. In a social
context in which sexual violence is normalized, a survivor giving an account of their
experiences may be blamed for what happened and socially shunned by friends, family, and
their community more broadly. An underling politely telling their boss that it is not okay to
make racist jokes in front of the new client risks professional repercussions as they are
perceived as difficult and a trouble-maker. And sometimes these jeopardizations manifest as
physical threats: a woman at a bar rejecting the advances of an interested man has reason to
fear he will respond with violence, as has happened far too often. Even the seemingly
innocuous statement that “Black Lives Matter” carries with it real risks of violent responses
from those who perceive it as a threat against the current social hierarchy.

Risky speech calls into question or directs critical attention to an aspect of our current
practices in a way that pushes back against those in power, and in doing so opens the
speaker to risk. It is not merely the content of the speech that makes it risky, but to borrow
Austin’s term, “the total speech act situation” -- the intersection of content, context, and the
speaker’s identity (Austin 1962). It is far more precarious for members of marginalized
communities to speak out about oppression than for members of privileged communities to
do so. Social privilege can mitigate the risks -- what counts as risky for one speaker may not be for another -- though it rarely banishes them altogether.

Both accusations and reports can be instances of risky speech, though neither always are. When a person reports on their experiences of sexism, Islamophobia, ableism, or any other of the myriad forms of bigotry in our world, they are engaging in risky speech -- even when their purpose is merely to call attention to what they or others have experienced. Reporting on each of these forms of injustice calls existing power dynamics into question, and makes the speaker vulnerable to professional, social, and physical risks. Moving to hold someone accountable for such unjust behavior is similarly risky.

Risky speech is liable to be taken up as an accusation, even when the speaker intended it as -- and ostensibly met the conditions to issue -- a report. This is because risky speech calls attention to, or calls into question, social hierarchies and so may be met with incredulity. It might not be apparent to an audience that the reported violation is a type of violation that can or does occur. For example, when whiteness and white privilege are the social norm, the mechanisms that maintain this hierarchy are often invisible to those who benefit from them. Subtle instances of racism may simply not be “on the radar” of white and other racially privileged people; these sorts of actions aren’t easily epistemically accessible to them. So a report about experiencing racial microaggressions may be met with disbelief that such things happen, let alone that it happened in this instance. Taking up this account as an accusation legitimizes this skepticism by shunting the speech into a category where skepticism is not only warranted, but appropriate.

Shunting risky into the category of accusations distorts what the speaker is able to do with their speech. It blocks off the speaker’s ability to issue reports -- to issue speech about
their experiences that calls for belief, and in doing so to enter into a normative relation of trust with the audience. This kind of disempowerment is an instance of injustice – both discursive and otherwise. More than that, this kind of response structures the material responses to risky reports in particularly harmful ways. In order to show this, I turn to an example of risky speech being met with the objective response.

III. Survivors’ speech

Sexual violence is both widespread and normalized. One out of every six women and one out of every thirty-three men has been the victim of attempted or completed rape (RAINN n.d.). Yet sexual violence is typically treated as a background condition of daily life. People are taught to avoid being victimized, rather than not to commit sexual assault – in much the same way as people are taught to plan for the weather and other ‘natural’ phenomena. As such, survivors’ accounts are risky – speaking out about their experiences is a way to draw attention to the issue and push back against this normalization. It is well recognized that survivor accounts are frequently dismissed – sometimes bluntly, by dismissing sexual violence as something unworthy of time or attention, and sometimes more subtly, by casting the survivor as untrustworthy, overemotional, or confused. Increasingly, though, there is

---

3 I am going to follow standard practice and use ‘survivor’ to refer to a person who has been subject to sexual violence. I recognize, though, that not everyone with such experiences self-identifies this way – sometimes because they do not consider their experience to be egregious enough to warrant the term ‘survivor’ and sometimes because ‘survivor’ seems to imply that the experience is in the past, whereas their trauma moves with them through life. I also recognize that this is, to an extent, begging the question – using this term presupposes that the speaker really has been sexually assaulted. This last part is deliberate; given the widespread dismissal of experiences of sexual violence here I opt to err on the side of belief.
recognition that these moves are impermissible – sexual violence *does* matter and those coming forward ought to at least be given the benefit of the doubt.

Well-meaning people sometimes struggle with how to respond to these accounts. On the surface, it seems like the objective response is an exemplary response to survivor testimony – it recognizes accounts of sexual violence as being about a serious harm and it moves to attend to the survivor’s speech. Indeed, survivors often receive this response from friends, family, co-workers, fellow students, teachers, and, especially, online commenters responding to their experiences. This is well illustrated by the comment at the beginning of this chapter. This was posted in response to a woman who wrote an op-ed about being raped by a fellow student and her university’s lack of response:

> It’s appalling the way her case has been handled thus far. Still it’s important to note that the young man in question is innocent until proven guilty. I want to see her case go to trial and be judged on its own merits, but we must not jump to conclusions about the player’s guilt… I’m just trying to put some common sense into people who drag others name through the mud online before all the facts have been heard. Now the fact that she did a rape kit and is willing to prosecute speaks very strongly in her favor. But the people calling him a rapist, before he has been convicted, are just as bad as the people who call her a liar based on her alcohol-level and clothing. We should always help the victim, but we as a society cannot judge nor punish unless a person is convicted beyond a reasonable doubt. (Comment on Marans 2016)

This is a paradigmatic example of the objective response. As such, it is a clear move away from traditional forms of dismissal of survivor testimony. It takes the issue of sexual violence seriously and is even expressly attempting to place the speaker on the same credibility level as the person she has identified as her assailant. Yet, this kind of response ultimately ends up serving as yet another way to dismiss survivors’ speech.

First, as a default response to survivor testimony this presumes that such accounts are issued as accusations. But this is not always the case. Survivors of sexual violence speak
about their experiences to friends, family, and support groups. This can be a way to gain recognition of what has happened, explain changes in behavior, seek support, recover from trauma, or find membership in a community of people with similar experiences. Even when survivors come forward to workplace or campus administrators or to the public at large, it can be as a way to raise awareness of an existing problem with the climate at the institution or as an appeal to institute programming to prevent such violence from continuing. Often, the survivor deliberately chooses not to name their assailant. In all sorts of contexts, the survivor may not be issuing an accusation – indeed, when the assailant is not even named, it is hard to imagine how the speech could serve as a way to initiate holding that person to account.

Giving the objective response to survivor testimony can distort what the survivor is able to do with their speech. This response defaults to taking the survivor as issuing an accusation. By doing this, it places the survivor into the normative role of an accuser, rather than a reporter. Each of these roles has different carries with it different orientations towards others, benefits, and responsibilities. Being taken as an accuser blocks off the possibility for the speaker to successfully issue a report. Their speech is shunted into the category of accusation, rather than report. This prevents the survivor from being taken as issuing the invitation to trust in their first-personal experience that is constitutive of a report. In a social world where accounts of sexual violence are often dismissed or disbelieved, this invitation to trust is significant. Taking up that trust – both in the survivor themself and in the reality of their experiences – is a significant normative move. Trusting the survivor is a way to solidify their status within the moral community, while failure to trust reaffirms their
position on the margins. By blocking off their ability to successfully make this move, the objective response limits survivor’s agency.

Second, how a survivor’s account is taken up can impact whether the university, workplace, or community implements interim steps to secure members’ safety and wellbeing. If the survivor is taken as issuing a report, then the appropriate stance is to default to believe the content of what they have said. In light of this default belief, it is not only reasonable, but required to take interim steps to secure community and survivor safety and wellbeing. At universities, this can include connecting the survivor with support systems, instituting sexual assault prevention programming, and any number of preemptive steps. Similarly, employers have an obligation to attend to those they have some reason to believe are creating a hostile workspace. When survivors’ speech is treated as a report, this gives employers just such a reason. While these interim steps may be revisited or changed depending on the findings of an investigation, believing survivors allows these processes to get started. Of course, preventative measures such as data collection and preventative programming technically could be implemented at any time. But if administrators do not yet believe that something has happened, then it is less likely that these proactive measures will even show up to the officials as steps one might take. Without the default belief generated by a report that an assault took place, these steps are less likely to present to officials as pressing or needed interventions. When investigations can drag on for months or years, these interim steps can be meaningful safeguards for survivor and community wellbeing.

If, on the other hand, survivors’ speech is taken as an accusation, then the appropriate stance is to hold off on believing their account. Given this, there is little reason to take any interim steps to secure community and survivor safety and wellbeing; there is no reason to
believe there is anything to secure it against. Not only are interim steps then unnecessary, they are actually uncalled for and perhaps a waste of time and resources. When taking survivors’ speech as an accusation, it only makes sense to take steps of this sort after an investigation has taken place and corroborating evidence, if any exists, has been found to lend credence to the accusation. This means that in cases where an investigation drags on, nothing will be done until a conclusion has been reached, and then only if the accusation is found to be true.

Third, if survivor testimony is taken as an accusation this can impede the very possibility of ever believing what they have said. When survivors are taken as leveling accusations, the proper response is to remain skeptical about what they have said, to hold off on belief. This means that that testimony has already been cast as insufficient to warrant belief on its own. Very often there is no physical evidence that can definitively prove sexual harassment or assault took place. When physical evidence exists, many times it can be interpreted in different ways – bruises can be taken as part of rough though consensual sex, sexually explicit messages can be characterized as part of consensual relationship, etc. In these cases, and when there is no physical evidence at all, the survivor’s testimony is the only thing that exists to back up claims of sexual violence. When an investigation in response to an accusation seeks corroborating evidence to support the accusation, and the only evidence available is the survivor’s account, that evidence is already taken as suspect. The survivor’s words are not easily recruited to serve to support their accusation when their testimony already not believed.

Based on this, there is often little reason for the audience to ultimately believe survivors’ accounts. Instead, the performative structure of accusation leads to the audience siding with
the accused. It is noteworthy that the identified assailant is rarely subjected to such
skepticism – their speech is typically treated as a report, rather than an accusation, and they
are believed unless there is reason to reconsider this belief. Put bluntly: the identified
assailant’s speech is believed by default, whereas the survivor must labor under an (often
impossible) burden of proof in order to receive belief. This means that what originally
looked like an impartial and objective move, in practice amounts to prioritizing the speech of
the person identified as the assailant. That is, the objective response is not, in reality,
objective at all.4

This difficulty in recruiting the survivor’s speech to lend credence to their accusation
means that investigations are unlikely to find in support of the survivor. Not only are interim
steps to secure community and survivor wellbeing unlikely to be taken, but it is unlikely that
anything will be done at all. Though on the surface it appears that treating survivors’ speech
as an accusation is an ideal way to seriously attend to it, in practice this uptake skews the
pragmatic response towards disbelieving the survivor and leaving the status quo unchecked.

Mandatory reporting under Title IX can exacerbate this.5 Under Title IX’s mandatory
reporting requirement, when a university representative is made aware of gender-based

---

4 This practice of prioritizing the testimony of the accused over the accuser seems to stem
from our judicial practices. But those practices explicitly acknowledge that the practice isn’t
neutral – it’s ostensibly weighted in favor of protecting the defendant. Hence the mantra that
it is better for 10 guilty people to go free rather than 1 innocent person to go to jail. I am not
taking a stand one way or another on what our judicial practices ought to be. However, I do
urge us to recognize the distinction between judicial practices and those we utilize in daily
life.

5 Title IX is, overwhelmingly, an important and positive tool for fighting back against sexual
and gender-based violence. The mandatory reporting requirement is intended to prevent
university officials from using their discretion about what is “worthy” of university
involvement or attention. One of the positive aspects of mandatory reporting is that it
facilitates better data collection about sexual violence at the university, and hopefully enables
violence, harassment, or discrimination occurring at or in relation to the university they must disclose it to a designated official (usually the Title IX coordinator) within 24 hours. This means that when a survivor talks about their experiences to a professor, coach, RA, or other university representative, that person is required to pass along what they have said. Often, survivors have confided this person in order to seek recognition, support, or accommodation. They are reporting their experience. Yet once mandatory reporting comes into play, officials work to determine whether a violation has occurred, and if so, what the appropriate punishment ought to be. ‘Susan,’ a survivor at the University of Arkansas, described her experience after being sexually assaulted by a resident advisor while incapacitated by a medical incident:

Confused and unsure of what to do following the alleged assault, Susan said she went the next day to a trusted older person she worked with to ask her advice. The friend stopped her in the middle of her story. “Before I finished even explaining it, she said, ‘What you've explained to me is a sexual assault, and I am a mandated reporter under Title IX, so you'll have to go through the Title IX process,’” Susan said. “Right then and there, it went from getting advice to me being in the Title IX process.” (Koon 2016)

Within the mandatory reporting system, the survivor’s speech is taken up as an accusation. (The ‘reporting’ of mandatory reporting identifies what the university representative must do – not what the survivor is taken to be doing.) This is often done by well-meaning officials who are working to properly respond to sexual violence on campus. But by taking up the universities to implement targeted and productive interventions in, say, teams or clubs that have a track record of repeated offenses. Here, I am identifying one hazard of this well-intentioned practice.

Many universities are earnestly working to implement Title IX requirements in an empowering way. As schools are working to develop best practices for this, there is wide variation in how these requirements are implemented. Some schools have adopted policies that navigate these pitfalls well, but the practice discussed here is one of the dominant implementations.
survivor’s account as an accusation, the survivor is propelled into a process they didn’t choose to enter into. What started off as an invitation to trust in the survivor and their experiences results in the survivor being forced to speak and take on an impossible burden of proof, only to be ultimately dismissed.

Once mandatory reporting has activated, the survivor is faced with a choice: participate in this process, or not. If they do not engage with this process, they risk losing control of their narrative – they won’t know what information is being passed along, what steps are being taken, or who is involved. This can be extraordinarily disempowering and can compound the trauma of the initial assault. Retaining some involvement and control in the process is a way to regain a sense of agency; this is especially important after the trauma of sexual violence. Yet if a survivor engages with the mandatory reporting process, this too can be traumatic and disempowering.

Survivors can feel that they do not really have a choice except to participate in the mandatory reporting process – their account is, as Rachel Ann McKinney terms it, *extracted speech* (McKinney 2016). Extracted speech occurs when a person is forced to speak. Sometimes this takes place in official ways, as when a reticent witness is compelled to testify in a trial. Other times it can be more informal. A well-intentioned friend might publically prompt another to bring up a topic they had no intention of talking about. The friend might feel socially compelled to speak; now that the topic has been broached it’s too awkward to remain silent. Both of these forms of extracted speech are potentially unpleasant – after all, the speaker is being compelled – yet neither are necessarily oppressive. Sometimes, though, a speaker may be compelled to produce speech that contributes to their own oppression or disempowerment – what McKinney terms *subordinating* extracted speech. This can happen
when police manipulate suspects into false confessions, or when an abusive partner compels their spouse to share degrading stories about themself.

Within mandatory reporting, survivors can experience their speech as compelled. Participating in the investigation is a way to retain a degree of control over the process, and remaining silent feels like a further erosion of the survivor’s agency. Yet being compelled to speak is also an infringement on their agency. Susan said, “The meeting basically looked like, ‘So you’re in this, and you have the chance to have a voice. If you want out, you can go out, but the process is going to go on without you and your name may get dragged through the mud in the process, so if you want to stand up for yourself, you need to be there,’” (Koon 2016). It is a classic double bind, in which one’s options are narrowed and each option comes with significance drawbacks (Frye 1983). More than simply being compelled to speak, survivor’s speech is taken up in such a way that furthers their disempowerment.

When a survivor is compelled (explicitly or subtly) to talk about their assault under Title IX, their speech is extracted. When their speech is taken up as an accusation, the survivor’s speech is subject to the burden of proof that accompanies accusations. But this burden of proof is often not easily met, especially when the survivor’s own testimony doesn’t count as supporting evidence. Even when universities adhere to the preponderance of evidence

---

7 Consider how the word ‘alleged’ is deployed in these proceedings. In order to maintain ‘impartiality’ within the investigation, university officials typically refer to the ‘alleged’ assault. Many people have critiqued the way this word is disproportionately used in relation to experiences of sexual violence, while left out when discussing accounts of other crimes or wrongdoing. What hasn’t been discussed, though, is how this treats the survivors’ speech. ‘To allege’ means to claim that a person has done something wrong, without evidence that this happened. When officials refer to the ‘alleged’ sexual assault, they imply that the survivor’s speech doesn’t count as support that the assault took place. The point is not just that survivor’s speech is not taken as sufficient evidence to prove that it occurred, but rather that their speech does not count as evidence at all. In contrast, referring to the ‘reported’
standard, which requires only a 50.1% chance that the reported assailant is responsible, survivors are at a disadvantage: when their speech is taken as an accusation it is, properly, not granted default belief, while the identified assailant’s speech is taken as a report and is, again properly, granted default belief. Taking up the speech as an accusation creates a situation where it is structurally reasonable to not believe the survivor. Thus, a survivor can first be compelled to speak and then on the basis of that speech be placed in a position of increased vulnerability. The survivor is now a participant in their own disempowerment.

Susan says of her experience within the mandatory reporting system: “It was, to this day, the worst experience of my life. I say it’s worse than the assault. It truly was” (Koon 2016). The RA who Susan reported as sexually assaulting her was found ‘not responsible’ by the university.

The objective response facilitates giving uptake to intended reports as accusations. As such, it is a mechanism of discursive injustice – it enacts skewed uptake of survivors’ speech acts in such a way that furthers their disempowerment. More than that, it structures the material steps taken (or not) in response to survivors’ speech. What seems like a helpful process to identify and combat sexual violence on campus instead ends up furthering the survivor’s disempowerment, lack of agency, and ostracization. Even when performed by those with good intentions, the objective response as the default reaction to survivors’ speech plays a potent role in maintaining existing systems of sexual violence.

incident at least acknowledges that the survivor has done something, they have reported it. And that there is reason to believe that it happened specifically because they reported it and so have vouchsafed that they experienced this violation.
IV. Objectivity, masking, and legitimizing

Risky speech is often met with resistance. Melissa Harris-Perry identifies “Prove it!” as one of the common responses given when people recount experiences of racism (Harris-Perry 2011). One way to make sense of this response is as a form of epistemic violence: it casts the speaker as an incompetent knower, as someone who is confused or mistaken about the world around them and their own experiences. Those recounting experiences of racism are most often people of color. Calling for proof, in this view, is a way to make clear that members of this group are considered suspect epistemic agents; that the audience deems them to have low credibility (Dotson 2011; Fricker 2007).

While it is certainly the case that some of these calls for proof are due to racist skepticism about the speaker’s credibility, this does not yet fully account for the resistance that identifications of racism tend to receive. A white person pointing out a friend’s subtly racist comment frequently results in denials and skepticism. In part, this because no one likes to think of themself as the sort of person who gets such important things wrong – it is far more comfortable to deny that the problem exists than to acknowledge one’s own harmful behavior.

But also present is a misunderstanding of what such identifications are doing. Pointing out a friend’s biased actions is often intended as a report – a way to alert them to what they are doing with the hope that they will consider and alter their behavior in the future. This kind of move calls for trust: trust in the second-personal relationship between the speaker and audience as well as trust that the speaker is in a position to properly identify the egregious behavior, typically in virtue of their life experiences or expertise in the area. Yet
rather than taking up the report as such, the friend often takes it as an accusation, as if it were a call for holding to account and for punishment of some sort.

The response signifies that if racism exists, the only reason to talk about it is in order to blame someone. This forestalls productive conversations about how to make things better. But even more fundamentally, it shifts the conversation from being about the experience or enactment of racism to being about the person identified as behaving this way and whether and how they are accountable. It makes the conversation about the person with racist behavior, rather than about the person subject to such behavior or about how that behavior is part of systemic oppression. Taking up reports of racism as accusations shifts the focus of the conversation by centering the wrongdoer, often at the expense of shunting the person who experienced their racist behavior to the margins.

Responding with “what’s your proof?” is a way to deflect the attribution of wrongdoing. Often well-intentioned people do this, even without realizing it. Whether intended or not, though, it is a way to maintain the audience’s position within the social hierarchy (DiAngelo 2011). When risky speech is routinely taken up as an accusation rather than as a report, this serves to obscure and (by doing so) uphold the status quo. It casts statements that resist or call out the status quo, such as identifications of racism or sexism, as requiring support for their believability. As Harris-Perry writes, “In a nation with the racial history of the United States I am baffled by the idea that non-racism would be the presumption and that it is racial bias which must be proved beyond reasonable doubt. … If anything, racial bias, not racial innocence is the better presumption” (Harris-Perry 2011).

This is not to say that the objective response is wholly separate from acts of epistemic violence. If someone’s speech is (unjustly) deemed lacking in credibility, one way to obscure
this is to treat the speech act as an accusation: not believing an accusation is proper, whereas not believing someone’s report is improper. Taking up their speech as an accusation, rather than as the intended report, can obscure this low credibility attribution. The objective response can serve as a masking move: it obscures the other acts of injustice, and does so in a way that makes it harder to identify where things have gone wrong.

As a masking move, the objective response gives plausible deniability to the audience. After all, the objective response is a response to the speech, so they cannot be accused of simply dismissing it. And, if the speech act is an accusation, it is correct to hold off on belief until more evidence is produced so the person giving this response cannot be accused of treating these accusations differently from any others. This response gives the veneer of correctness. It legitimizes the biased behavior, and further re-entrenches the initial impulse to dismiss the risky speech. Credibility attributions are entwined with the category of speech act the utterance is shunted into and given uptake as, and this in turn influences the appropriate responses and next steps.

The objective response serves as a legitimizing move. Remaining skeptical of the content of an accusation is not simply permissible; it is appropriate. Similarly, asking for evidence in light of an accusation is the proper next step. By taking up claims of racism as accusations, this legitimizes both lack of belief and the increased burden of proof placed on the speaker. That is, in doing these things, the speaker is ostensibly behaving appropriately and even laudably, even while these moves undermine the speaker. The objective response, and the role it plays in distorting risky reports as accusations, is a subtle way to dismiss risky speech while maintaining the veneer of acting correctly.
The objective response is lauded because it is seen as an impartial and measured form of response. It appears neutral, and as if the audience is responsibly removed from the situation. This form of response is connected to what Sally Haslanger has described as objectifying reason (Haslanger 2012). Objectifying reason is a projective epistemic and practical stance. Here, an observer picks up on certain social truths but – crucially – fails to note how these truths have been socially created and produced, and the role they themself have played in this production. Instead, the observer takes them as natural. The veneer of objectivity obscures reality while at the same time (falsely) assuring the observer of their objective understanding of reality.

The objective response is connected to this assumed objectivity. The objective response serves as a mechanism to mask and legitimize mistrust of the speaker. More than this, it obscures role of the audience in producing the speaker as a marginalized subject. The veneer of objectivity (falsely) assures the audience that they are behaving impartially and are merely responding to the speaker in appropriate ways. But by beginning with mistrust for the speaker, and then using the objective response to legitimize this failure to trust, the audience is retrospectively constituting the speaker as an appropriate target for mistrust.

V. Conclusion

Accusations and reports are performatively distinct, but easily mistaken for each other. I’ve argued that risky reports are especially likely to be taken up as accusations, which both distorts what the speaker is trying to do with their speech and can ultimately undermine the
speaker. The objective response masks and legitimizes this undermining move. One possible response to this trend is to err on the side of caution: to take on the default assumption that speakers issuing risky speech are intending to issue reports unless otherwise specified, and to take up their speech as such. This strikes me as a flawed response. This default uptake would make it more difficult for survivors and others issuing risky speech to receive uptake when they intend to issue accusations. And even though accusations can put survivors’ accounts in a precarious position, they are also an important discursive tool. Accusations are a way to initiate ameliorative projects – they are a mechanism by which we can get retribution, or reparations, or an apology. I’m wary of limiting the agency of survivors and others issuing risky speech by making it harder to use this discursive tool.

This leaves us with the messy task of deciphering when to take up risky speech as a report and when to give uptake to it as an accusation. Both accusations and reports are important discursive tools. Inviting trust and belief in oneself and in one’s experiences is a significant move, and is especially so when injustices such as racism and sexual violence specifically erode this kind of belief and full inclusion within the moral community. At the same time, identifying wrongdoing and calling to hold the perpetrator to account is also a crucial tool. Especially for people who are already marginalized, identifying this kind of injustice and calling to hold the perpetrator to account are potent moves. Without this, we have few ways to correct for egregious behavior. We need to maintain speakers’ ability to issue both reports and accusations without having a strong default reaction to taking up risky speech as one or the other. We need to work to develop better sensibilities for distinguishing reports from accusations. This is our difficult task going forward.
We use speech in myriad ways to signal, construct, and negotiate our identity in our daily lives. Sometimes this is explicit: saying, “As a woman, I have all sorts of thoughts on the CDC’s 2016 recommendation about alcohol,” calls into salience the speaker’s gender and invites others to treat what she says as coming from a particular, gendered, perspective. Other times this identity signaling is done through subtle methods: speaking a dialect can solidify that one “really” belongs to a particular community, as does using words or phrases in particular ways that have specific meaning only to members of an ingroup. Similarly, we use language to call into salience aspects of other people’s identities. Sometimes this functions as an invitation for them to take on a particular normative position in virtue of that aspect of their identity, say as an expert witness or as a person who has experience being on Welfare and can offer advice to others. Other times this is done to people; discursive moves call to salience an aspect of their identity, and it’s pragmatically difficult, if not impossible, for the person targeted to extricate themself from the position that’s been thrust upon them.

Recently, there’s been a surge of work on how identity impacts what people can do with language: how things like race or gender play a role in whether a speaker is recognized as a knower and thus a credible speaker, whether they are able to secure uptake as performing an intended speech act, or whether on the basis of their identity they are subject to threats of violence as a response to their speech. Here, I am interested in the other side of the coin:
how we use language to do things with identity and the real impact this has, both good and bad, on our social interactions.

I offer a pragmatic account of what I term discursive signaling. Discursive signaling is when we use linguistic moves to call into salience particular aspects of our identity in a given context. This kind of calling to salience helps to structure how we and those around us understand and navigate that interaction. As Christine Korsgaard writes, “If I call out your name, I make you stop in your tracks…Now you cannot proceed as before. Oh, you can proceed alright, but not as you did before. For now, if you walk on, you will be ignoring me and slighting me. It will probably be difficult for you, and you will have to muster a certain active resistance, a sense of rebellion. But why should you have to rebel against me? It is because I am a law to you… I have given you a reason to stop.” (Korsgaard 1996, 140).

Discursive signaling creates reasons for the audience to treat a person in a particular way, in virtue of the aspect of their identity that was called to salience. More than merely making that aspect of the person’s identity salient, signaling calls it to salience by calling on others to recognize it as such. It is a particular second-personal move. While the audience can disregard this call, they now have to ignore it or dismiss it – neither of which are the same as if nothing had happened. And if they disregard the call to salience, they have disregarded something that the speaker has called on them to attend to rather than simply disregarding a feature of the world. As such, discursive signaling makes particular second-personal claims on its audience.

Discursive signaling also plays a role in structuring the identity of the person it is about; it can be a form of interpellation. Interpellation is “the process by which people are constituted as persons with particular locations in social normative space through vocative
calls and the acknowledgements they demand” (Kukla and Lance 2009, 181). First theorized by Althusser, interpellation is the process of subject formation in normative space. As Althusser explained it, interpellation is the process by which a hail places a person in a particular social role and demands that they give enacted recognition of their subject location. Hails like, “Hey there, friend!” locate the target as a particular kind of subject, in this case a friend to the speaker, and call for the person to recognize *themselves* as already inhabiting this role. Interpellation occurs throughout daily life, and this repetition helps to solidify the subject roles we are placed in. Interpellation is the process by which our subjectivity is enacted upon us from the outside, which we then internalize and take up ourselves as if that’s who we were all along.

Discursive signaling doesn’t merely call attention to an aspect of someone’s identity; it can also help to produce that part of their identity. When a teacher addresses their classroom, “Hello students! Are you ready to learn today?” they interpellate the children in the room as students. By calling on them specifically as students, the teacher calls on the children to recognize themselves as inhabiting the role of ‘student’ and all the attendant norms and obligations that go along with being a student in that classroom and school. The teacher has called to salience their student-ness, which calls on each child to recognize themselves as a student and to do this in virtue of the teacher having called on them to do so.

I begin by discussing the complexities of social identity. Then, I present the pragmatics of discursive signaling as a kind of speech act. In the next two sections I explore how we signal aspects of our own identity and how we do this to others. In each, I explore both the positive and negative aspects of this kind of discursive move, and the real implications it has on how we understand and interact with others. Slurs, I argue, are particularly powerful tools
for discursive signaling. In the third section I explore two different ways we do this with our own identity: signaling directed at other ingroup members, as a way to constitute a collective “we,” and signaling directed at outgroup members, to call on outsiders to attend to an aspect of our identity they may have otherwise failed to regard as significant in that moment. In the fourth section I explore the ways we call aspects of others’ identities into salience: through hermeneutic force, in which we take persons to be inhabiting particular social positions and understand or interpret their actions and perspective in light of this standpoint, and through interpellation, by which we hail others as inhabiting a particular social position and demand that they recognize themselves as such.

I. Situated and salient identities

We all have complex and intersectional identities. We inhabit various communities that shape our worldview and sense of self, and others perceive and treat us in ways that shape our experiences and how we interpret the world. The boundaries of these communities are often neither rigid nor clearly delineated; people's identities are complex and multi-faceted, with membership in many communities, ranging from communities organized around race or ethnicity, gender, sexuality, family, friendships, social interests, professional pursuits, or complex intersections of the above. Sometimes these are separate groups, while other times one group is nested within another, or they contain areas of salient overlap. Here I introduce three aspects socially situated identities: 1) complex social identities, 2) intersectional identities, and 3) “worlds” of which we are a part, and then discuss the notion of salient aspects of our identities.
Kristie Dotson and Marita Gilbert introduce the notion of complex social identities, saying, “Complex social identities simply refer to the fact we are all simultaneously part of multiple communities that can influence how we are “read” or understood in social space. … Complex social identities, then, refers to identities that are readable in social landscapes according to our (real or perceived) membership in multiple communities whose membership has hermeneutic impact” (Dotson and Gilbert 2014, 875-876). That is, each of us inhabits multiple communities, and people respond to us in important ways depending on which community we’re perceived as belonging to in a particular context or interaction. How people perceive and respond to us then shapes our experiences in that interaction. This can have important ramifications, from whether or not we’re granted community membership, to the opportunities we’re offered, to whether people properly recognize and attend to our material needs.

In addition to inhabiting multiple communities, we also have intersectional identities. Whereas some approaches to identity take aspects of identity as separate, so that, for example, gender and race are seen as distinct identities and systems of oppression, intersectionality attends to the interconnections between gender, racial, and other social structures (Crenshaw 1997). In this way, women of color don’t merely exist under a ‘double burden’ of racial and gendered oppression, but rather their experiences of race structure their experiences of gender, and vice versa. Whereas identity politics such as the civil rights movement or Black Lives Matter, the feminist movement, or the gay rights movement have largely focused on the experiences of the dominant group members within each movement (black men, white women, and affluent gay white men, respectively), intersectionality calls for recognition of the diverse experiences of members of social groups. Even these social
groups aren’t homogenous, and other aspects of our identity may and often do structure our experiences within those communities – both by community members and by outsiders. For example, I’m woman, but I’m not just a woman, I’m also a philosopher, I’m white, I’m a New Englander, I’m a feminist, I’m straight, I’m able bodied, I’m a scifi geek, I’m socioeconomically privileged, I’m a friend and a family member, as well as all sorts of other things. These aspects of my identity interlace and intersect to structure my particular experiences in complex and context dependent ways. So my experience of womanhood is structured by the other aspects of my identity, but just how these aspects of identity come into play will vary depending on the situation, who I am with, and all sorts of other contextually dependent factors.

María Lugones introduced the notion of *worlds*: worlds are built around our shared experiences with others and the worldviews that arise from shared aspects of our identities (Lugones 1987). However, worlds are more than a group of individuals coming together and sharing a particular perspective: they are tied to particular social and material locations in which these people interact. Worlds can be large or small, and there can be worlds embedded within each other. So the particular world of a lower-class mestizo Latina will be importantly different from the world of an upper-class white Latina, though both may be part of a the more far-reaching Latinx world. We all inhabit multiple worlds, and often the set of worlds an individual inhabits can wildly dissimilar from one another. So a person who travels between dominant hearing culture and deaf communities may find that these worlds are more disparate than a person who travels between the worlds constructed around being an Oregonian and a New Englander.
Within each world, we are a particular aspect of our self. We switch between these worlds all the time, some people more often than others. Lugones identifies this process as *world-traveling*. We shift from one aspect of identity in a particular world to another in a different world. Not only are our identities complex, they are sometimes seemingly contradictory: as Lugones puts it, we are “ambiguous” beings with multi-imaged selves” (Lugones 1987, 13). Our selves aren’t necessarily unified or cohesive – who we are in one social location can seem wildly contradictory to or incompatible with who we are in another. The norms we adhere to, the values we espouse, and how we perceive and interact with the world can differ widely depending on community context. That is, aspects of our selves and the worlds we inhabit are often dissonant with one another.

While we can’t completely set aside all these complex identities, communities, and worlds we inhabit, often particular aspects of identity can be especially salient in a given context. That is, a particular aspect of our identity may be called into salience, which then structures how we navigate a situation and how others view us. Sometimes what is pertinent is determined by structural features, such as the context or setting: college roommates may interact with each other primarily as friends or roommates when out and about on campus, but if they take a course together, once they enter the classroom and the class begins their identity as students is brought into salience and is generally the facet of identity that structures their experience in that setting. Their identity as friends, or as a sibling to someone, or as a particularly skilled musician falls into the background. (Note that here the setting itself structures which aspect of their identity is salient, rather than a discursive signal doing this work. If they ignore their status as students and continue to interact as roommates while in the classroom they may be behaving oddly or inappropriately, but they haven’t necessarily
resisted a second-personal claim made upon them.) Other times, which aspect of identity is salient is governed by others: we are interpellated into particular social roles based on how others view and treat us. Sometimes this is relatively innocuous: a medical doctor having lunch with a friend may find herself unwittingly placed into the role of medical professional when the friend starts asking for clinical advice. Other times this takes a more pernicious turn: women and people of color at professional events may expect their profession to be at the forefront, but instead find that their gender and/or race is taken as most salient by the people they interact with. This can manifest in misevaluating the person’s credibility and authority, mistaking their interest in the topic as being casual or superficial, or wrongly characterizing them as being in a subservient nonprofessional role at the event. That is, just which aspect of identity is taken up as salient doesn’t just impact the ease or comfort with which we navigate a particular situation, it also has real material consequences.

Often, just which aspect of identity ought to be salient is underdetermined. While in the context of a professional event, professional expertise in that area ought properly be taken as at least one of the salient parts of a participant’s identity. (There are things that can alter the context – such as police showing up looking for a witness to a crime.) Often, though, it’s unclear just which “hat” a person ought to be wearing or taken as wearing – there are multiple positions from which they might be speaking, or which others might properly treat them as inhabiting. In a casual discussion on reproductive justice, I might reasonably offer contributions specifically from my perspective as a feminist, as philosopher, or from my own personal experiences. While sometimes these aspects of identity can be reconciled with one another (my perspective as a philosopher is deeply entwined with feminist thought), sometimes just which world we are speaking from will yield distinctly different perspectives.
This is particularly the case when we inhabit dissimilar worlds and have dissonant aspects of self – in these cases, just which aspect of our self is made salient can have a profound impact on how things go.

This process in which particular aspects of our identity are called into salience in a given context or situation is what I term *identity signaling*. Our identities are complex and socially situated, and within this we convey information about our identities in a multitude of ways. Sometimes this process of identity signaling happens without our conscious control, such as when skin pigmentation is taken to convey information about racial identity. It can also happen intentionally, as when a person wears particular clothing intended to convey information about their gender identity. *Discursive signaling*, then, is when we use linguistic moves to do this work. In the following, I’m specifically interested in how we use discursive signaling to do things with our own identity, especially as a means to resolve the indeterminacy of which perspective we’re speaking from or as a method to resist improper interpellation.

II. Discursive Signaling

*Discursive signaling*, in which we use discursive moves to make salient particular aspects of identity, is a particular subset of identity signaling. Sometimes we signal our identity explicitly, with statements like: “As a feminist, I care deeply about reproductive justice,” or “I’d like to purchase a ticket with the veteran’s discount; I’m a veteran.” Often, though, we use more subtle discursive practices to signal our identity. Sometimes this happens at the level of semantic content: which topics we choose to bring up and discuss. When I frequently talk about issues relating to sexual violence and feminism, others might start to
pick up on my identity as a feminist. Other times, it can happen through how we speak: accents and dialects are taken to convey a whole host of information about the speaker. While often accent and dialect are outside the control of the speaker, sometimes a person may consciously and intentionally switch between accents or dialects depending on the context – when in a formal setting, a southern accent or speaking Black English can be a hazard because of the stereotypes associated with the speakers of each, whereas when with friends or family using these may be ways to signal community membership. Similarly, code-switching, when bilingual speakers mix and switch between two languages when in conversation together, is often used as mechanism to signal ethnic or cultural identity together.

Austin famously introduced the idea that we do more with speech than convey meaning; speech is also an action that can effect changes in normative space (Austin 1962). Austin argued that speech acts are significant beyond their truth or falsity: they’re significant because of what they do. Importantly, speech acts are complex, and a single utterance can do multiple things at once. Exercitives, one kind of speech act, call forth the normative conventions to which we are bound. That is, they don’t merely call attention to existing rules; exercitives make it the case that these rules are the ones to which we are bound in a particular situation. So, enacting laws or creating rules are paradigmatic instances of exercitives. More subtly, using racial slurs in a public setting can enact norms that license or endorse racist behavior (McGowan 2012).

Discursive signaling is a form of exercitive. It calls into salience particular aspects of a person’s identity in a given situation. This calling into salience is the paradigmatic pragmatic force of these kinds of speech acts – it makes it the case that this aspect of the person’s identity
is important in this particular moment. More than that, it’s bringing into being the particular lens through which the speaker and the audience ought to be interpreting the subject, either explicitly or tacitly, in the discursive context.

Discursive signaling can take various forms, and each form enacts a particular and distinct illocutionary force:

*First-personal discursive signaling.* Example: “I’d like a veteran’s ticket; I’m a veteran.” Here, a speaker makes salient an aspect of their own identity. The force of the speech act is reflexive. In speaking, the speaker is calling attention to an aspect of themselves that will structure both how the speaker navigates the context and how others ought properly “read” and respond to the speaker.

*Second-personal discursive signaling.* Example: “You’re a doctor, what do you think?” This form of discursive signaling is essentially interpellative (Kukla and Lance 2009, Chapter 8). In an interaction with another person, this speech act makes salient an aspect of that person’s identity. They are placed into a particular normative category, which comes with distinctive norms of behavior, and relations to others. Here, which aspect of their identity is made salient is outside the person’s control. While it’s possible to resist these kinds of interpellative moves, it’s pragmatically complicated to do so, especially if the person does in fact have the aspect of identity being called forth.

*Third-personal discursive signaling.* Example: “Next time we see her, we should get her perspective on the crisis in Flint; I bet as a black woman she’ll have lots of thoughts on it.” This speech act makes salient an aspect of the target’s identity when that person is
being talked about rather than spoken to. This kind of signaling has a particular hermeneutic force: Rather than directly placing the individual into a particular normative space, it does so at a level removed. This speech act structures how others view, interpret, and respond to the subject though this may not be at a level that is directly accessible to the subject themself.

In all these cases, it’s not the specific grammar of the speech act that constitutes it as first-, second-, or third- personal, but rather what the speech act is doing, where and how the pragmatic force is directed. It’s well recognized that the surface grammar of a speech act need not dictate the type of act being performed. When a teacher tells their students, “Please turn in your papers on Friday,” the teacher isn’t really issuing a request, regardless of how it’s phrased. Rather, this is a command, and any student who treats it as a request, as optional and if performed as deserving of gratitude, has done something wrong. In the same way, the kind of signaling a speaker is performing isn’t determined by the surface grammar of the utterance. We can perform reflexive discursive signaling without first-personal utterances, such as by telling inside jokes to call forth and establish a speaker’s insider status. Similarly, second-personal signaling need not have a second-personal grammatical structure, as with, “Gosh, it sure would be nice if we had an expert around to tell us how to do this,” said leadingly to the room at large but obviously calling on a particular person known to have expert knowledge to take up that role. Instead, discursive conventions, such as context, tone, and conversational norms, govern what’s being done.
Reflexive discursive signaling makes salient aspects of the speaker’s identity in a specific situation. This is the only form of discursive signaling that can be intentionally undertaken by the person whose identity is being called into salience, and as such is one of the ways by which we can do things with our own identity in social spaces. Reflexive discursive signaling is an exercitive that enacts a norm in which the speaker makes salient aspects of their identity, which then structures how both the speaker and the audience interpret what’s going on. Yet this only scratches the surface of the pragmatic contours and the importance of these kinds of moves. Reflexive discursive signaling can further break down as signaling to ingroup members or signaling to outgroup members. Each case calls for the audience to take up the speech act through their shared recognition as inhabiting a particular category.

Kukla and Lance (2009) introduced the distinction between speech acts that are structurally agent neutral in their outputs, and so can be taken up by anyone, from those that are structurally agent specific, and so are intended for a particular person or persons to take up. To this, Kukla and I added the further category of community specific speech (Herbert and Kukla 2017). Community specific speech, we say, “functions, as a matter of its pragmatic structure, to call for uptake from members of a community qua community members, and not just qua individuals who have the shared property of belonging to a community” (Herbert and Kukla 2017, 580). Community specific speech calls for an enacted recognition of a “we” constituted by the audience and speaker’s membership in a shared community.
We see this kind of speech all the time. In technical jargon and acronyms, dog whistle politics, or even in playful references and insider jokes that only ingroup members can understand. The point here isn’t simply that there may be semantic content available only to ingroup members, but rather that part of the structure of this kind of speech is to bring to the forefront the shared community membership. When a group of teenagers speak to each other by referencing a series of inside jokes, part of what they are doing is calling to into salience their shared membership in that particular friend group. Though the jokes aren’t intended for any individual member of the group *qua* individual, they are intended for members of the group *qua* friend group members. By offering these jokes, and by laughing in response, the speaker and audience both enact their membership in the particular group.

Community specific speech calls for and enacts a shared recognition of group membership, a collective “we,” whereas reflexive discursive signaling calls for others to attend to the speaker’s membership in a group—sometimes those others are also part of that group and sometimes they are not. In cases where the audience is part of the group being called to salience, this is what I term *ingroup* signaling. This co-travels with community specific speech; one can’t occur without the other. In the case of inside jokes shared amongst a friend group, part of what’s going on is the signal of the speaker’s identity—they are speaking *as a member of that friend group* and are calling on others to treat them as such.

In other cases, signaling occurs in contexts where other group members aren’t present; in these cases of *outgroup* signaling, the speaker is calling on the audience to recognize their status *as* outgroup members. This kind of speech act is neither agent neutral nor community specific; it inhabits another, though more attenuated, category. This signal calls for others to recognize that the speaker inhabits an identity *unlike theirs*. Here the point is that the speaker
is calling into salience a facet of their identity that others may be ignoring, be ignorant of, or simply not understanding the importance of in the given context. Calling this aspect of their identity to salience is a way to flag that the speaker is acting from a normative position outside of those occupied by others in the context. Both, “move out of the way, I’m a doctor!” and “as a Latina, here’s what I think about the discipline of philosophy” call for others to attend to the speaker’s skills, status, or knowledge that may otherwise have been missed. Whereas ingroup signaling calls for shared recognition amongst the speaker and their audience, outgroup signaling calls into salience their differences.

Discursive signaling calls into salience a particular aspect of the speaker’s identity. In doing this, it seeks to bring about real normative changes. 1) It constitutes that particular aspect of the speaker’s identity as salient in the interaction, which impacts how others hold and interact with that person. 2) It constructs the normative space as being a place for that kind of person. And, 3) it enables recognition of complex and intersectional identities. However, these changes aren’t assured, and sometimes things go quite wrong.

Reflexive discursive signaling constitutes a particular aspect of the speaker’s identity as being salient in the interaction, which impacts how others hold and interact with that person. Sometimes this is overt, as when a speaker calls into salience an aspect of their identity in order to gain entry to a space or to benefits, as with “I’m a pilot; I should be let into the pilot’s lounge” or “I’d like a student’s ticket, please.” Other times, this signaling structures a series of ongoing interactions with the speaker, as with, “I’m not going to talk about work today; today is all about family.” This kind of statement calls into salience the speaker’s status as a family member, and calls for others to set aside work-related topics and issues and instead interact with them qua their family member identity.
Just which aspect of their identity is called into salience structures how others are called on to interpret the speaker’s perspective and knowledge claims. “Well, back when I was at Harvard . . .” dropped casually into a conversation functions as a call for authoritative status; the speaker is calling forth the cachet of an elite educational background, which, presumably, credentials them on a whole host of topics. On the other hand, signals like “I’m a doctor!” call for recognition of their authority in a particular field or subject. Often, the point of this kind of signaling is to call recognition to the speaker’s situated knowledge – that they are offering a perspective from their status as inhabiting a particular social location. To other ingroup members this works to call for shared recognition of knowledge claims – the sense that we know this. Outgroup signaling calls for recognition of the importance of the speaker’s outsider perspective. This is especially the case in contexts where the presumption is that the participants are homogenous or share a universal or group perspective. This kind of signaling can call for a reevaluation of the speaker’s perspective and utterances – that even if they are expressing something foreign or dissimilar to others’ experiences they oughtn’t be simply dismissed precisely because they are coming from an alternate social location which may be inaccessible to others.

Of course, sometimes this can take a negative turn: we internalize damaging norms and oppressive self-conceptions, which we then draw on and exert in interactions with others. Someone who has been persistently exposed to the relentless body-shaming and fatphobia in our society may say, “I’m fat; I’m worthless” and genuinely expect others to take on this perspective of them in their interactions. Signaling can be self-undermining in a dazzling variety of ways, from devaluing our own knowledge with statements like, “I’m just a girl, so you might know better, but here’s my opinion . . .” to using non-reclaimed slurs to call to
salience our (self-perceived) deficient character and moral status. Doing this both affirms these oppressive norms, as well as calling to salience the speaker’s subject location within them.

Reflexive discursive signaling can also serve to constitute the space in which the interaction is taking place as being a place for that kind of person. In spaces that are overwhelmingly white, calling explicit attention to one’s non-white race can help to construct that space as being a place for people of color. Sometimes signaling works as an act of resistance, as when women and people of color signal their credentials in professional contexts where they are being trivialized or ignored. Yet other times constructing the space as being for or open to particular identities can be an act of solidarity or welcome – it can be a way to showcase inclusivity and openness. And, importantly, it can be an act of joy or pride – we don’t only call forth aspects of our identity when it’s necessary to do so, but also when we’re proud of that part of ourselves, when that is something we find joy in. This is the project of revalorized and reclaimed identities – to place the identity at the forefront and hold it as valuable. Reclaimed slurs, in which the derogatory term has been recast by the targeted group as being either a neutral description or a positive affirmation, is a paradigmatic example of this. Reclaimed slurs are a way to call forth the aspect of identity targeted by the slur, but to upend valence of the term and showcase pride and joy in that aspect of identity.

Finally, this signaling enables recognition of complex and intersectional identities. By calling forth an aspect of our identity that may not have been immediately apparent to others, we call for an enacted recognition of the complexity of identity – that identity isn’t something that can be wholly read or sussed out by others. Even when we’re with people who know us
well, signaling can alert them to the process by which we don and take off “hats” and perspectives. The point isn’t merely that we all engage in this kind of project of centering certain aspects of self, but rather that discursive signaling make this explicit and publicly recognizable. That is, signaling makes this process a public social fact, and others are called on to take up not only what we have done but also that we have done it; not only which identity we’ve made salient, but also that we’ve engaged in a project of making salient an aspect of our identity. This moves away from the notion of unified and rigid selves and toward the notion of selves as playful, complex, and socially and contextually situated.

At its core, reflexive discursive signaling is a way of enacting the speaker’s agency – agency over their own identity, and setting norms for how others ought to hold, respond, and treat them in that context in virtue of their identity. But no matter how clear and directed a speaker is, they can’t be assured that things will go as intended; as with all discursive interactions, signaling relies on the audience to give the proper uptake to the speech act. And sometimes, this doesn’t happen. Reflexive discursive signaling can go wrong in several ways: 1) the signal is ignored, 2) the speaker’s identity membership is rejected, 3) the relevance of that identity is denied, or 4) the audience takes the signal in an unexpected, and unwelcome, direction. Each of these is, at least some of the time, tied to forms of identity-based injustice.

Sometimes a person’s move to call to salience an aspect of their identity fails to receive uptake from the audience. It’s not that no one heard or saw the signal; rather, the signal is simply ignored. A woman of color at a professional conference may seek to exert her credentials in response to being dismissed by other participants, only to be greeted by silence and a continuation of the dismissive behavior. Langton introduced the notion of
illocutionary disablement to describe instances where a speaker is unable to bring about the intended illocutionary act with their speech; they are unable to receive uptake as performing that act at all. Importantly, illocutionary disablement is norm-governed; it’s not simply that sometimes people make a mistake about what a speaker was trying to do, but rather that social norms, usually connected to identity, come into play to structure the way an audience takes up the speech act. In this way, social norms that sexualize non-consent and cast women as overtly against sex but secretly willing, can come into play to block the proper uptake of women’s sexual refusals. Similarly, when a speaker inhabiting an already marginalized identity uses speech to signal their expertise, that marginalized identity may structure the audience’s response to prevent the signal from being properly taken up. Put simply, one aspect of the speaker’s identity may, due to oppressive social norms, block their ability to call into salience another aspect of their identity.

In other instances, a speaker’s signal may be properly taken up, but then called into question. The same speaker trying to exert her credentials may, instead of silence, be greeted by incredulity. The audience may either reject the speaker’s claimed membership in the community or they may deny that that group membership is relevant in the moment. That is, the speaker may express her credentials and expertise, only for the audience, either explicitly or implicitly, to question whether she really warrants that expert status despite her objective claim to such (Fricker 2007). Once again, aspects of her identity are coming into play to

---

8 In the Chapter 4, ‘The Pragmatics of Reclamation Projects,’ I introduce a distinction between illocutionary silencing, in which the utterance has no performative force, and illocutionary distortion, in which the force of the speech act is something other than that intended by the speaker. This erasure of a reflexive signal is an instance of illocutionary silencing.
undercut her credibility to exert another aspect of her identity; attempts to push back against this are often read as emotional hysteria rather than authoritative statements (Kukla 2012). Similarly, light skinned Latinx may face this when ingroup signaling to other Latinx, only to find that their lighter skin is taken to undercut their claim to Latinx community membership.

Or, perhaps, the audience may deny that the claimed identity is salient in the situation. If the speaker is working to make salient a marginalized aspect of their identity, then mainstream audiences may push back against the perceived intrusion of that identity into the social space. Here, the identity being asserted is, itself, what blocks the intended effects from taking place.

Finally, sometimes a speaker is successful at calling to salience an aspect of their identity, only for the audience to demonstrate their ignorance of that identity (Dotson 2011). This can manifest as a misunderstanding, say, by mistaking what it means to say that one is a first generation American, which can lead the interaction in unforeseen and unexpected directions. Or, this kind of ignorance may manifest through demonstrated failure to properly understand group membership in that community, say through tokenizing the speaker and generalizing the speaker’s perspective to represent the entire community, with statements like, “You think this, so now we know what all black people think about it. Great!!” Rather than successfully pushing back against social norms that marginalize the identity, this response to the identity signal ends up re-entrenching the identity’s marginalized status.

At its best, reflexive discursive signaling can be a powerful way to enact the speaker’s agency over their identity and how they are read by others in a given situation, either by constituting ingroup community membership or by calling others’ attention to an erstwhile
unattended standpoint. However, these effects are not assured and reflexive discursive signaling doesn’t come without hazards for the speaker.

IV. Signaling others’ identities

Not only do we use language to call into salience aspects of our own identities, we do this with other people’s identities as well. Whereas reflexive discursive signaling is a way for the speaker to enact agency over and regarding their own identity, signaling others’ identities is the opposite: a way for the speaker to enact agency over and about another person’s identity in a given social interaction. These sorts of signals have hermeneutic force: they set norms that structure how an audience understands and interprets the person whose identity is being signaled. When that person is present in the interaction, this also functions as an instance of interpellation: it places the person in a certain social position, and calls for them to recognize and take up this role.

Hermeneutic force structures how we “hold” or “read” people. Statements such as “Obama, the first black president, spoke after the Charleston massacre,” puts both Obama’s race and status as the president of the United States at the forefront, and calls on others to interpret his actions through his status as both a black man and as a powerful authority figure. Hermeneutic force determines through which “lens” we view a person in a given situation; different lenses will pick up on different information and yield different understandings. Recognizing that someone is acting in their capacity as a parent may yield a substantially different reading of a confrontation between strangers in a parking lot; they may
be acting to protect their child’s safety, rather than acting from the position of a road rage-filled asshole.

Calling to salience an aspect of someone’s identity can happen with or without that person being present in the interaction. Most people will never interact directly with Obama, yet still interpret him and his actions on the basis of particular identity categories. Other times a person may be, or come to be, part of the interaction. They may slowly pick up on the ways in which they are being hermeneutically read, via the contextual clues in the interaction, or they may not. The point is that it’s not primarily about the person whose identity is being called to salience, but rather is about the other participants and how they are reading, and subsequently responding, to the individual on the basis of particular aspects of their identity.

When discursive signaling is used to call into salience an aspect of another’s identity, this works as a form of interpellation. It locates that aspect of identity as the subject role the target is already inhabiting, and calls on the target to take up and act in this role. Importantly, interpellation is difficult to resist; efforts to do so often end up reinforcing the very structure that was being resisted. Saying, “Dude, we’re not actually friends” still encodes the recognition that the hail was intended for that person and that they inhabit the status of being the sort of creature who could be recognized as the speaker’s friend. That is, even denials of interpellation often still reinforce the social structures the interpellative move was enacting.

Just as with reflexive discursive signaling, signaling other’s identities can be a mechanism to promote agency or it can be a mechanism to diminish or restrict this agency. At its best, signaling others’ identities can be a way to show value for others’ experiences, knowledge, or
agency. Utterances like, “You’re the pregnant one, it’s happening in and to your body; it’s your decision what happens next,” can be a way to show recognition for the real differences social situatedness has on decision-making and agency. Similarly, signaling others’ identities can be a way to call to salience their expertise in contexts where it may have gone unrecognized: “You know her from other contexts, but Deserea’s actually a bioethicist – next time you see her, you should ask her these questions about Zika.” Or sometimes it can be a way to celebrate an aspect of another’s: “Tonight we are here to honor Dr. Jimena for the role she has played in developing [exciting things] in molecular biology!”

In each of these cases, the call to take an aspect of their identity as salient can serve as something akin to an invitation. The person can take up the call from the signal or not, though now that it’s been extended as an option acting from that particular standpoint is more than a neutral possibility like all other possibilities. The speaker has expressed an interest in hearing from them *qua* that identity, but doing so is, hopefully, not being forced upon the person. The important point is that this kind of signaling, when done right, comes with open exit conditions. It’s possible for the pregnant person to say, “I’d rather not focus on that; I’d rather we make this decision together” or for the bioethicist to say, “I just taught three classes, I’m super burned out and want to just grab drinks and chill, but maybe we can chat about it some other time.” It may require practical negotiation to sort out under just what conditions this kind of signaling realistically leaves open the exit conditions: it’s likely that doing this in private carries with it less normative pressure than doing so in public, though it will depend on all kinds of contextual features. Sometimes these negotiations may take place beforehand, so that the speaker gains prior authorization before this kind of signaling: such as when Dr. Jimena agreed to the event in her honor as a molecular biologist,
therefore giving tacit approval for signaling this aspect of her identity. Presumably, if this were a role she did not want to take up, she could have declined the event in her honor.

Navigating and negotiating these exit conditions is tricky, and unfortunately not all instances of signaling other’s identity comes with such accessible options for opting out. Far too often, signaling others’ identities forces the person to take on that normative role. Sometimes this can occur from well-intentioned people, as when an overzealous partner tries to showcase their significant other’s accomplishments, much to that person’s horror: “You have experience with this, go on dear! Tell them! No? You don’t want to? Well, here, I’ll tell them for you.” Here, the exit conditions of the interpellation are increasingly less accessible to the person being targeted; they are being interpellated into this subject position whether they wish to be or not.

As I’ll discuss in greater detail in the following chapters, the use of slurs is a particularly powerful way to do this. The exit conditions from slurs and other unwanted signals of identity are pragmatically difficult to navigate. It’s hard to say “yes, that is a thing about me, but it’s not relevant here so please don’t take it into consideration.” Either this ends up unintentionally denigrating that aspect of identity or it just further calls it into the spotlight. When a marginalized aspect of identity has been called to salience, not wanting to focus on it is often read as dismissing or expressing shame in that aspect of identity; it’s much easier to set aside a privileged or neutral aspect of identity. Even if a person successfully navigates asking others not to attend to that aspect of their identity in the context, this request itself can serve to further call attention to that part of the speaker’s identity resulting in stereotype rebound.
At its best, signaling others’ identity can be a way to show value for the person’s agency and identity, yet this kind of signaling is risky precisely because it can so easily slip over into constraining the target’s agency.

V. Conclusion

We use speech to do things with identity all the time in our social world. These are ongoing practices to which we are all party. At their best, these practices of calling to salience aspects of identity can be ways to secure and support agency: they can expand the possibilities for entrance into social spaces and can encode those spaces as open and inclusive and can showcase value for different standpoints. At their worst, though, these practices can severely restrict agency, placing the target into a constrained social role or further re-entrenching the person’s marginalization. However, by attending to these practices, we can better position ourselves to promote what is valuable in them while attempting to safeguard against what is pernicious.
CHAPTER 3

DEROGATIVES: INSULTS, SLURS, AND SPEECH ACTS

Philosophers are increasingly recognizing and attending to the power of harmful words (“Slurs” 2013; Croom 2015; Maitra and McGowan 2012). But to the degree that the conceptual terrain of harmful words has been explored, it’s typically been framed in terms of slurs and insults, contrasted against one another. According to the rule of thumb used to differentiate them, an insult targets individuals based on behavior, whereas a slur targets people based on membership in a social group. Slurs especially have been the subject of a great deal of rich and fruitful work, and rightly so. Yet framing the issue of harmful terms in this way is unnecessarily limiting our understanding. This distinction between slurs and insults doesn’t helpfully account for the often murky line between group membership and action. Nor does it allow for the conceptual space to capture the rich variation amongst harmful words more generally.9

There is a great deal of variety amongst harmful words, both in terms of their history, the sorts of relations they call forth, the contexts in which they function powerfully, and the way they’re used and deployed in everyday speech. The terms ‘slut,’ ‘basic,’ and ‘Becky,’ for example, all target women on the basis of their gender and so (on some readings) might all be simply classified as slurs. Yet they have distinctly different meanings and effects, and are

9 Exceptions to this include Sarah-Jane Leslie’s focus on generics, Jennifer Saul’s work on dogwhistles, and Lynne Tirrell’s work on toxic speech. These works successfully move beyond the slurs v. insults paradigm for thinking about harmful words, though they generally leave the category of ‘slurs’ untouched.
used in quite different ways. ‘Slut’ is recognized as an offensive gendered slur for women and is taboo (even though it’s still frequently used), whereas the others are relatively commonplace and largely considered acceptable. All of these terms have at least the veneer of being warranted in virtue of a woman’s behaviors and activities – whether that’s sexual activity, a certain set of derided interests, or a kind of oblivious social climbing that comes at the expense of people of color. ‘Slut’ has a long and rich history behind it, whereas ‘basic’ and ‘Becky’ are quite new. While ‘slut’ can be directed at any woman, though with different effects depending on who its directed at, ‘basic’ and ‘Becky’ are specifically directed at white women. Finally, while ‘basic’ is used to disparage white women and their interests in a way that reinforces patriarchal power relations while at the same time reinforcing whiteness as a transparent default, ‘Becky’ is used by women of color to ridicule and call out the privilege of a certain kind of white femininity. Our existing conceptual framework for thinking about slurs simply can’t capture these sorts of differences, and even expanding the framework to include both slurs and insults doesn’t much help.

Framing the category of harmful terms solely in terms of slurs versus insults has two interrelated effects. First, it flattens the conceptual terrain -- setting up this dichotomy narrows the focus to paradigmatic insults, and paradigmatic slurs – primarily, those that are already recognized as such and carry a strong social taboo against their use regardless of context. This doesn’t leave much conceptual space for the subtle, complex, and contextual ways harmful words can operate in our daily lives. I’m interested in attending to words that are socially acceptable, yet still do the derogating and dehumanizing work of slurs, words that subordinate their targets when deployed within certain communities but that don’t have the same significance elsewhere, and words that draw on more recently-enacted hierarchies.
and oppressive structures than do typically recognized slurs. Second, this focus on paradigmatic slurs as contrasted with insults has helped to blur together the semantics and pragmatics of the terms. By focusing on only the typical cases, cases in which insulting terms are deployed in insults and slurs are deployed in speech acts that derogate, this facilitates focusing on *only* these cases and not on the complicated and messy ways slurs can be deployed in a variety of speech acts. Shifting this focus reveals both the broad array of words that can be recruited for derogatory speech and the complicated life of slurs outside of their typical deployment.

In the following I aim to expand the conceptual taxonomy of harmful words. First, I'll identify the generally agreed-upon features of slurs. Second, I'll offer expanded conceptual machinery for thinking about slurs. I'll argue that there is a multidimensional continuum of terms, which I term derogatives, that share some or all of the features of paradigmatic slurs. Just how a derogative is located on this multidimensional continuum impacts how it functions in the world. Third, I'll explore the connections between derogatives and speech acts. Derogatives are paradigmatically deployed in derogatory speech acts. But not all derogatory speech acts involve derogative terms, and derogative terms are deployed in many speech acts other than derogatory ones. Finally, based on this expanded conception of harmful words, I'll focus on the various ways derogatives can enter into and go out of our lexicon: via linguistic shift in which derogatives become obsolete, the “euphemism treadmill” in which new words come to be slurs because of stigma about the associated group, or through intentional reclamation projects.
I. Features of slurs

There’s been much ongoing debate about what, if anything, differentiates slurs from other categories of terms. However, five features are generally agreed upon: slurs mark out membership in a social group, they essentialize some (real or perceived) characteristic or characteristics of that group, they dehumanize or derogate members of the group, they draw on and are connected to a history of oppression of the group, and they set up norms for the treatment of members of that group.

First, slurs identify members of a social group – often on the basis of race, gender, sexuality, gender identity, nationality, religion, or disability. A slur picks out a person or persons in virtue of this group membership. A social group is a group of people who are bound together by shared life experiences – by choice, socially imposed grouping, or some combination. Particularly important is that in order to count as a social group, there needs to be some sense in which this group membership plays a constitutive role in meaningfully creating the person’s identity (i.e., it is in being treated as a group member that one comes to be a group member). Slurs pick out members of a group who have importantly similar life experiences in virtue of a particular identity feature. At the same time, being the target of a particular slur can be one of the mechanisms by which the boundaries of the social group are constructed and reinforced: shared experience of being the target of the slur can be an important part of the shared life experiences that bind together the social group and helps create the group identity. The application of slurs can also help to create and maintain the contours of this group identity (Kukla 2016; Swanson 2016). So, slurs both target social groups and help to construct and reinforce the contours of the group. By establishing and
maintaining these group contours, slurs also help to enforce an *us v. them* dichotomy. Those within the group marked by the slur are thought of as importantly different from those outside the group. More than this, by calling into salience social group membership, slurs help to enforce intergroup thinking, in which people understand the world in terms of group membership rather than in terms of interactions amongst individual persons.

Second, slurs are socially embedded, in that they draw on and call into salience a history of oppression of the group (Tirrell 2012). Slurs don’t come out of nowhere. They are often the product of a history of stigma attached to the associated group. More than that, slurs specifically mark out those with diminished social power in some way – those who are relegated to the margins of society or who are the subjects of systemic oppression. Slurs are hooked into these histories and systems of injustice, both making that entrenched oppression of the group salient when the slur is deployed, and reinforcing and perpetuating it. As Swanson puts it, slurs elicit or cue particular ideologies, and at the same time using a slur makes the ideology to which it is connected more acceptable (Swanson 2016). Slurs don’t pick out social groups imbued with power and privilege – or at least, they don’t pick out members of those groups *qua* the identity aspects that grant their power and privilege.

Third, slurs essentialize the social group in question. They identify (real or perceived) characteristics or sets of characteristics that are presumed to mark out the group as a distinct social kind (Leslie 2014). These essentialized characteristics are taken to be deep, immutable features of the group that differentiate it from others. These features are presumed to be natural and are often grounded in either supposed biology (skin color, reproductive capacities, or cognitive capabilities) or innate character (laziness, emotionality, or untrustworthiness). These presumed-intrinsic characteristics are then taken to have particular
moral implications: if someone’s skin color and innate character mark them as fundamentally lazy, for example, then they are best served by having others order and direct their lives. People who do this are merely fulfilling their moral obligations to members of the group (or so the slave-owning rationale went).

Part of what it means for a group to be essentialized is that group members’ actions, motivations, and capabilities are most readily understood within the context of the essentialized characteristics. If a black man is essentialized as having “bad character” then any apparent criminal activity on his part can be viewed as enacting that bad character, rather than, say, as the result of racial profiling that leads to black men being subject to high levels of police scrutiny and any ambiguous actions being more likely to be read as criminal. Anything that falls outside these essentialized characteristics may be disregarded or recontextualized to fit within the established framework. When group members’ actions are taken to conform with these expectations, this helps to reinforce the belief that the essentialized characteristic really is natural and immutable.

Fourth, slurs help to establish norms for interactions with members of that group. Based on the way the group is essentialized and the historical treatment of the group, the presence of a slur can call to salience particular normative relations towards members of the group. If a slur marks out members of the group as having bad character, then members of the group ought properly be treated with suspicion and wariness. If the slur calls to salience a system in which members of the group are thought to be flawed epistemic agents, then they ought not to be listened to or have their preferences taken into consideration. If members of the group are thought to be ruled by their animalistic nature, then the slur sets up relations in which others ought properly “help” the group members by ordering and directing their lives for
them. That is, the slur doesn’t simply call to salience a history of these kinds of behaviors, it also sets them up as the normatively appropriate moves to make.

Just how this works is a source of contention. Tirrell terms this feature “action engendering within a context” meaning that the presence of a slur calls for and authorizes particular actions depending on how and when it is deployed. Camp proposes that slurs work by licensing a particularly derogatory perspective of members of the targeted group, which then shades and influences audience members’ interactions with group members. Either way, slurs serve as tools to help enact norms for how members of the group ought to be treated, which can have profound effects outside of the particular interaction in which the slur is deployed.

Finally, slurs aren’t neutral descriptors. They derogate, dehumanize, or disparage members of the group they refer to. While there’s much ongoing debate about just how slurs do this derogatory work (via the stereotypes they convey through implicature, the inferential role they play, the hidden generic of the slurs, the disgust they express, the perspectives they license, or the history they draw upon) there is widespread agreement that slurs do something over and above more acceptable ways of referring to the social group – and that, however it’s cashed out, this extra work is bad for the targets. This is typically taken as the central feature of slurs, though it’s built up out of the other defining features.

These five features entwine and mutually support each other. The normative relations instantiated by the slur typically stem from the history of how the group is essentialized, which in turn structures the contours of how the group is dehumanized. These interactions then reinforce the shared life experiences of members of the group, further entrenching it as
a social group. While there’s debate about just how, exactly, these features of slurs work, there’s general agreement these roughly mark off the category of slurs.

II. Derogatives: Expanding the taxonomy of harmful words

So much for what’s generally agreed upon about slurs. In addition to these generally accepted features of slurs there are several others that are either tacitly or explicitly held up as essential: that the slur is taboo in contrast to a neutral counterpart for the group, that it has “derogatory autonomy” and conveys derogation regardless of how it’s deployed, that the term functions as a slur across all of society, and that it draws on a deeply entrenched history of oppression against a robust social group. Yet there are many words that fit the preceding five features of slurs without meeting these conditions. It would be troubling to presume that none of our currently acceptable words are harmful, or that we need to have first recognized a term as taboo for it to be harmful. In addition, there’s a slow but growing recognition of something like community-specific slurs – in which subcultures have their own slurs that are deeply dehumanizing to those targeted within the community but which are unintelligible to outsiders. Or contextual slurs, like ‘basic,’ that gain their power based on how the term is deployed and in what context. Similarly, “feminist,” depending on context, might function as a slur, an insult, a description, or a term of approbation. Finally, some slurs are an outgrowth of recently enacted systems or oppression, or target “thin” social groups that may not be as robustly constituted as social groups organized around race or gender.
In the following, I address the contested features of slurs and argue that for each feature we ought to adopt a continuum along which a term may fall. This will help to map out the conceptual terrain of a class of terms I term derogatives. Derogative terms encompass paradigmatic slurs, which often fall at the extreme ends of this multidimensional continuum. But derogatives also include terms with more complex locations, terms that essentialize their targets without being taboo, or terms that don’t elicit such deeply entrenched systems of oppression. This expanded taxonomy of harmful terms will equip us to better understand the rich variations amongst slurs and other derogatives and distinct harms such terms can help enact.

Taboo v. acceptable

Paradigmatic slurs are taboo. This is, largely, how we identify terms as slurs. These terms are prohibited from use in formal contexts, are contested when they’re included in print or on television or the radio, and using them is generally frowned upon. Anderson and Lepore have advanced the view that what makes a slur harmful is that using it violates a taboo (Anderson and Lepore 2013). Slurs are often described in contrast to what philosophers standardly identify as a ‘neutral counterpart.’ A ‘neutral counterpart’ is the accepted label used to refer to the group. But this sets up a false dichotomy: that accepted terms are neutral and that harmful slurs are taboo. But not all socially acceptable terms are neutral. Many terms that are recognized as slurs today were previously widely accepted, yet even then they essentialized and derogated members of the social group. Think about terms for sexual orientation: medicalized terms, like ‘homosexual’ were long considered the acceptable way to refer to the group, yet this isn’t a neutral term – it calls forth and reinforces a particular
ideology in which non-normative sexuality is medicalized and pathologized (Peters 2014).

This acceptable term helped to normalize and legitimize the institutionalizing, stigmatizing, and dehumanizing of gay folks. So as a general point, philosophers ought to move away from the phrase ‘neutral counterpart’ and instead adopt the convention used within psychology of referring to the currently socially acceptable term for the group as the ‘category label,’ without making any presumptions about the supposed neutrality of the term (Carnaghi and Maass 2007).

However, this practice has particular relevance for thinking about taboos and derogatives. Being able to identify a ‘neutral counterpart’ is used as a sorting mechanism for whether or not a term is a slur. Terms without a corresponding counterpart often aren’t flagged as slurs see: Bolinger 2017). Yet if a term has not been marked as taboo, there simply may not have been a push to develop a less-harmful term for the group. At least sometimes, our currently acceptable terms may be derogatives. It could well be the case that the terms at play in daily life are degrading and essentializing and that they license certain kinds of normative interactions with their targets, but that they are so normalized that there hasn’t been a push to develop another way to refer to the group marked out by the term. We should leave open the possibility that some of the terms in daily use are derogatives, and in the course of time we may retrospectively recognize them as such – as has happened before.

If we accept the possibility that derogatives need not be taboo, and may in fact pervade our daily life, then we’re in a position to recognize words such as “crazy” or “thug” as derogatives, and to better understand the derogating power of such words. Derogatives can exist on a continuum from fully taboo (the n-word), to generally acceptable (crazy), to
somewhere in the middle (*slut*). Recognizing this spectrum allows for pinpointing the *different* work that taboo and acceptable derogatives can do.

Acceptable derogatives can powerfully perpetuate and normalize the derogatory view of the targeted group. Acceptable derogatives are positioned to be used frequently and off-handedly. The lack of taboo means the term can be deployed in a much wider array of contexts, and including in formal or official speech which adds further authority to the perspective or ideology to which the derogative is connected. Moreover, because it’s not marked as violating a norm, this works happens under the radar, as it were. People are less likely to note the use of the derogative and the work it’s doing may not be picked up on as easily. Acceptable derogatives can contribute to a kind of hermeneutic injustice – the lack of taboo makes it much harder for those targeted by it to pinpoint and get widespread recognition of the harm of the term, and efforts to do so are often thought of as petty.

Once a term is marked as taboo, its covert work is limited. The contexts in which the derogative can acceptably be deployed are narrowed. Going against this taboo, then, has specific force – it signals a willingness to violate that taboo, to endorse the perspective linked to the term, and to take on social costs such norm violation might incur (Anderson and Lepore 2013; Camp 2013; Waldron 2010). This pinpoints this aspect of the harm of taboo derogatives to breaking the norm against their use, rather than as a feature of derogatives generally. The power of a derogative is *different* depending on whether or not it is taboo, or how strong the taboo against its use it. Whether taboo derogatives are more harmful than acceptable derogatives, or the other way around, is an open (and likely contextual) question.

Genuinely making space for the possibility of acceptable derogatives is difficult and uncomfortable work. It means that we can’t rely on intuitions about which words count as
derogatives, and we can’t rely on our established practices to easily pick them out for us. Rather, we have to take seriously the possibility that our normal and everyday vocabulary might be doing essentializing and derogating work against particular groups of people, and that our language might make us unwittingly complicit in perpetuating their harmful treatment.

*Thick social groups v. thin social groups*

Slurs target people on the basis of membership in a social group. This a foundational aspect of slurs, and is what differentiates slurs from things like insults. The paradigmatic examples of slurs draw on well-recognized social groups, targeting people on the basis of such things as race, nationality, gender, religion, or sexual orientation. These social groups are robustly constituted: they are well recognized as social groups, and members of the group typically recognize shared life experiences on this basis of that identity. However, not all social groups are so robustly constituted. Social groups exist along a continuum, from those that profoundly structure group members’ lives and self-understandings to those that are more diffuse and have a shallower impact on the group members. ‘English-speaker’ may constitute a thin social group, one of which most people are largely unaware until they’re in a situation with non-English speakers. Yet the language we speak and think in has been shown to impact how we perceive the world and to shape our interactions (Boroditsky 2009).

Derogatives can target either ‘thick’ or ‘thin’ social groups.

This continuum means that the line between slurs and insults, originally so stark, may get murky at times. Whether or not ‘stupid’ targets a social group is, I think, up for debate. On the one hand, it is typically used as an insult, targeting a person’s bad decision making or
academic performance. But it also could be taken to target the group of persons who have low intelligence or certain mental disabilities. Given that this group is subject to ridicule, stigma, and diminished life prospects, there’s good reason to categorize ‘stupid’ as a derogative – though a derogative targeting a group that rarely recognizes itself as such.

Another example: ‘snowflake’ has recently come into vogue as the preferred insult against liberal activists. ‘Snowflake’ identifies the targets as being too easily offended or riled up over inconsequential issues, thus setting up a norm for easily dismissing the concerns of those people, and it conveys contempt for its targets. Yet it’s far from obvious that liberal activists constitute a robust social group. They might, but if so it’s a particular kind of social group, one that needn’t carry over into all, or even most, aspects of their lives. One could easily be a dedicated liberal activist on weekends, with work colleagues, family, and friends none the wiser. This ability to compartmentalize how the group identity permeates one’s life also means that the norms authorized by the derogative term needn’t spill over into other areas of life. A liberal activist may be able to critique others’ suggestions or behavior at work without those actions being viewed through the lens of being a ‘delicate snowflake,’ whereas race or gender identities are more likely to travel with a person through different contexts. Recognizing what sort of social group a term refers to, and why, allows for more fine-grained distinctions amongst derogatives.

Deep history of oppression v. some disadvantage

Some derogatives are robustly socially embedded, drawing on a deep and extensive history of oppression against the targeted group. Other derogatives, however, elicit weaker systems of oppression or even merely disadvantage. This could be because the group
targeted by the term has only recently become the subject of discrimination, and so the mechanisms of discrimination haven’t had time to become as widespread or as socially embedded as they have for other systems of discrimination against other groups. Or, it could be that the group in question is subject to social disadvantage or stigma that, while real, doesn’t fit the criteria of oppression. ‘Bro’ might be such a word. ‘Bro’ is used to mark a distinct social group of young, typically white, “alpha male idiots … [A bro is] inarticulate, belligerent, talks about nothing but chicks and beer, drives a jacked up truck that’s plastered with stickers, has a rich dad that owns a dealership or construction business and constantly tells this to chicks at parties” (Urban Dictionary). ‘Bros’ are characterized as obnoxious and unintelligent, and marking someone as a ‘bro’ licenses treating them with a certain degree of condescension and avoidance. They face ridicule and might experience difficulty being taken seriously in non-stereotypically masculine settings (such as within art or theater communities). But while marking someone with this term can disadvantage them in some ways, ‘bros’ also experience pervasive privilege on the basis of race, gender, and socioeconomic status. The continuum of whether the derogative draws upon a deep history of oppression against the targeted group or merely upon a pattern of disadvantage specifically appeals to the accepted features of slurs laid out earlier. Marking out how the term falls along these various spectrums enables a far more precise understanding of what a term like ‘bro’ can do.

*Derogatory autonomy v. discursively contextual*

Finally, philosophers have largely taken ‘derogatory autonomy’ to be a defining feature of slurs (Anderson and Lepore 2013; Camp 2013; Jeshion 2013; Tirrell 2012). ‘Derogatory
autonomy’ means that the term conveys derogation or contempt regardless of the speaker’s intentions. A speaker who lovingly says, “You’re a beautiful [slur],” may intend to compliment the recipient, but regardless of intentions, they have also derogated and dehumanized them. Similarly, slurs are recognized to ‘leak’ of denials. A speaker who says, “Jackson is not a [slur]!” still seems to be implicated in the work of the slur even though they’re denying its application. Here, that seems to be because they’re leaving open that the slur *does* properly apply to some people, just not to Jackson. However, denials of slurs still ‘leak’ even when the speaker is denying their application universally: “There are no [slurs]!” still calls forth the derogatory perspective and, on Elisabeth Camp’s view, licenses this perspective (Camp 2013). One way to make sense of this, which I propose in the final chapter, is that the mere presence of the slur queues a set of pernicious and dehumanizing implicit associations, and importantly, it does this particularly powerfully in contrast to other terms. That is, slurs express, signal, or call forth derogatory representations of the targeted group, regardless of the speaker’s intentions or how the term is discursively embedded.

Yet again, I urge us to consider this to be one end of a spectrum along which derogatives can exist. Some derogatives derogate regardless of how they’re deployed, while others do so only in particular discursive contexts. ‘Gay’ today is largely recognized as a pure description of sexual orientation. But the general negative, “that’s so gay!” uses the term as a derogative – albeit in a roundabout way (which I will discuss below). Unsurprisingly, it’s been found that people who use ‘gay’ as a general way to express a negative evaluation (roughly as a synonym for “that’s bad”) have increased homophobic implicit associations – though those who use the term as a pure description do not (Nicolas and Skinner 2012). This gives some
reason to think that the mere presence of certain derogatives is sufficient to elicit pernicious implicit biases, but that others depend on the discursive context in which they’re deployed.

Connectedly, terms that are in the midst of reclamation may sometimes do this derogatory work, and sometimes not. ‘Queer’, the most successfully reclaimed word to date, is widely used within academia and LGBTQIA communities as a pure descriptor. Yet it still functions powerfully as a derogative when hurled at someone as part of a derogatory speech act. I take up this aspect of derogatives, that their force is sometimes dependent on the ways in which they’re discursively deployed, in a great deal more detail in the following section. For now, the point is that the degree of derogatory autonomy a term has ought to be recognized as one more aspect of derogatives that lies along a spectrum.

_Society-wide v. community-specific_

Discussions of slurs have typically focused on terms that, barring reclamation projects, have a stable meaning across a given society. Paradigmatic slurs derogate the associated group regardless of where or when the slur is deployed. However, presuming this holds for all derogatives has two problems. First, the contours of ‘society’ here aren’t well defined. The default presumption is either that the term holds across the nation or across all speakers of that language. Neither of these notions of ‘society’ are defended and both are arbitrary. Second, this rules out derogatives that are specific to particular communities or sub-groups. Yet the terms used within these contexts can be powerful. Those who can use the term properly can solidify their insider status, while those who misuse it mark themselves as outsiders to the sub-community (Herbert and Kukla 2017). We ought to recognize a continuum, from society-wide to community-specific derogatives. Sometimes these
community-specific derogatives are unintelligible to outsiders, while other times they elicit alternate or weakened stereotypes. However outsiders read them, the term still holds powerfully within the particular community.

‘Fuckboi’ is an excellent example of this. The term started off within black communities in the 2000s and has since migrated into more widespread use (Shepherd 2014). However, it still operates almost exclusively with a particular sub-community, that of young people. Much more so that other young-person slang, ‘fuckboi’ is nearly incomprehensible to those twenty-five or older. Despite what outsiders might think, the term doesn’t focus on the sexual practices of its target. Rather, ‘fuckboi’ identifies are particular sort of (generally, but not necessarily) young man who is, “loosely, a streetwear trend-humper who … [pursues this with] a certain blend of awkwardness and thirst, combined with lots of disposable income … that translates into a need to be associated with the culture it inhabits, while perhaps not actually being of said culture” (Shepherd 2014). “Generally, fuckbois all have the same motives: To occupy time and space with the sole output being ‘hype’” (Dakota 2015). More derisive than ‘bro,’ ‘fuckboi’ (like ‘basic’) identifies privilege even while denigrating those who have it.

_Homogenous v. intersectional_

The existing literature on slurs makes a tacit assumption that they impact the targeted group in a unified, homogenous way. There has been little focus on or recognition of the differing effects an individual term has on differently positioned members of the targeted group. Slurs are either taken to affect everyone equally, or, if there is variation, that this variation is due to a person’s subjective experience and individual psychology. Yet this
certainly isn’t the case for all derogatives -- and may not be the case for any. We need to attend to the ways in which derogatives differently impact members of the targeted group depending on other aspects of identity – the intersectional effects of derogatives.

This point was eloquently raised by a group of Black women responding to the SlutWalk movement’s efforts to reclaim the word ‘slut’: “As Black women, we do not have the privilege or the space to call ourselves ‘slut’ without validating the already historically entrenched ideology and recurring messages about what and who the Black woman is. We don’t have the privilege to play on destructive representations burned in our collective minds, on our bodies and souls for generations” (Anonymous 2011). As these authors recognized, derogatives targeting a group on the basis of one aspect of identity – in this case gender – have disparate effects on members of that group depending on other identity features such as race, class, age, or disability. The stereotypes the derogative elicits, the history it draws on, and the treatment it authorizes are often varied, and varied in established and predictable ways.

Rather than continuing to rely on the tacit assumption that derogatives impact group members in a homogenous way, we need to attend to the intersectional aspects of derogatives. This is particularly important as it is typically the experiences and effect on the relatively disempowered members of the group that go unnoticed. The SlutWalk prioritized the experiences of white women’s gender-based oppression at the expense of black women’s voices being heard. This is part of a pattern within identity politics in which the experiences and needs of the relatively privileged members of the group are centered. We ought to be careful not to presume that derogatives impact all group members in the same way, especially when doing so ignores those who may be especially vulnerable.
A more complete taxonomy of derogatives, then, recognizes each of the following rows as a continuum:

<table>
<thead>
<tr>
<th>Taboo</th>
<th>Acceptable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thick social group</td>
<td>Thin social group</td>
</tr>
<tr>
<td>Deep history of oppression</td>
<td>Some disadvantage</td>
</tr>
<tr>
<td>Derogatory autonomy</td>
<td>Discursively contextual</td>
</tr>
<tr>
<td>Society-wide</td>
<td>Community-specific</td>
</tr>
<tr>
<td>Homogenous</td>
<td>Intersectional</td>
</tr>
</tbody>
</table>

A derogative can fall anywhere on the continuum for each row. For some derogatives, how it places for one of these features will impact its placement for others. Derogatives with a great deal of derogatory autonomy, for example, are likely to draw on a deep history of oppression, be highly taboo, and to function across many contexts in society. However, acceptable derogatives may vary quite a bit more, as may community-specific derogatives. Each of these metrics impacts what a derogative does: how it functions in our social world and the ways by which it does this work.

Expanding the taxonomy of derogatives enables us to more precisely identify and understand the nuances of complicated terms. ‘Basic,’ for example, is recognized as a dismissive or demeaning descriptor, though it’s hardly taboo. It’s a gendered term, though it doesn’t refer to just any women, but rather is used to targeted white women or those thought to be acting like white women. It specifically marks out a kind of white womanhood that is thought to be uninteresting or lacking in individuality. ‘Basic’ women like pumpkin spice lattes, Ugg boots, and leggings (Brown 2014). This term perpetuates gendered power dynamics that trivialize and disparage the interests and activities of young women, and in this sense draws on a long history of gender-based oppression. But it also specifically locates whiteness as a baseline. ‘Basic’ appeals to a system of white privilege that treats whiteness as
both universal and lacking in individuality, and as invisible and unremarkable (DiAngelo 2011). In this sense, it diverges greatly from paradigmatic slurs. ‘Basic’ both denigrates white women on the basis of their gender while at the same time maintaining their racial privilege.

‘Becky’ is a derogative that identifies a group similar to that picked out by ‘basic,’ yet it does different work. This term again marks out a particular group of white women, though here it identifies a kind of social climbing that comes at the expense of people of color. A ‘Becky’ is likely to be unaware of or unresponsive to social justice issues; she’s not ‘woke.’ Instead she’s focused on achieving social prestige, and is likely to do this by using people of color as a tool to achieve her ends. She may date men of color in order to improve her social status, or she may make use of issues pertaining to people of color for her own benefit. This term frames the target’s behaviors in terms of this kind of self-centered and not-socially aware social climbing. And it certainly isn’t complimentary. But the term is used primarily by people of color – especially women of color. It’s not reinforcing an existing history of oppression, but rather is a way to highlight and resist exploitation of communities of color. While it fits some features of paradigmatic slurs, it sharply diverges from others. The term is highly community-specific and discursively contextual.

Derogatives that diverge from paradigmatic slurs, like ‘basic’ and ‘Becky,’ help bring to light the murky line between insults and slurs. Traditionally, philosophers have appealed to the distinction that slurs draw on social group membership to identify their targets, while insults are applied in virtue of the target’s behavior. But a term like ‘basic,’ and many other derogatives, could easily be read as being about both identity and behavior – it’s a term applied to particular people who do certain things. In this case, white women who wear leggings, drink pumpkin spice lattes, etc. A more apt distinction is: insults are applied in
response to the target’s behavior, while derogatives are applied in virtue of membership in a social group and are used to contextualize the target’s behavior. If someone behaves badly, a person might label them with an insult as a way of describing and condemning their bad behavior. But if the speaker applies the term on the basis of the target’s identity, and then uses it to as a filter to interpret their behavior, the term is being used as a derogative.

This has two important implications. First, a discursively-contextual term can function as either an insult or a derogative depending on context. Luvell Anderson has written about how ‘thug’ is a racialized term that has particular effects on black men (Anderson forthcoming). But it’s well known that ‘thug’ is applied to people besides black men: it’s generally applied to criminals, especially those who engage in violent behavior. One possibility is that when ‘thug’ is applied to others, it’s serving as a metaphor – a way of saying that these targets are as bad as black men. I’ll discuss this kind of indirect usage in detail below, but this seems unnecessarily convoluted in the case of ‘thug.’ More likely is that this term functions sometimes as an insult and sometimes as a slur. When it’s applied as the result of criminal behavior, ‘thug’ serves as an insult – all it’s doing is picking out and describing that criminal behavior, albeit in an informal and condemnatory way. But sometimes it’s applied to black men merely on the basis of their status as black men. Anderson, a black man, describes his experience during middle school of being viewed as a ‘thug’ while on his way to serve as a crossing guard. Nothing about his behavior would be reasonably described as criminal. Rather, his gender and racial identity were taken by a passing police officer as sufficient to warrant the label. This then served as a lens through which the officer viewed his behavior. Similarly, political protesters in Baltimore, who were predominantly black teens, were labeled by the mayor and other national figures as ‘thugs.’
Their race, age, and the fact that many of them were men served as the basis for use of the term. The media dismissed their politically-driven actions as rioting and criminal activity (Pitner 2015). In contrast, violence committed by predominantly white crowds is generally not labeled this way – even when there are high-levels of property damage (Stanek 2014).

Locating discursively-contextual derogatives can better enable us to pick up on when a term is being deployed as a derogative or as an insult.

Second, the distinction between insults and derogatives lends itself to abuse. Speakers can justify using ambiguous terms by pointing to the target’s behavior. Such is the case with ‘slut.’ ‘Slut’ ostensibly targets sexual promiscuity. Yet in a 2014 study of college women, Armstrong, et al, found that there was little link between actual sexual behavior and who was characterized as being a ‘slut’.

\begin{quote}
Woman 1: I was a virgin the first time I was called a slut.
Woman 2: I was too.
Other Woman: Really?
Woman 1: Yeah, because no one knew [I was really a virgin].
Woman 2: They all thought I slept with people. That’s what my volleyball coach said to all my friends, that I was the one that was going to be causing trouble when I get older…
First Author: What were they responding to?
Woman 1: Like, if masturbation were to come up . . . I wouldn’t be afraid to talk about it. I think people got the wrong idea from that.
Woman 2: In high school, they called me a cocktease. I didn’t do anything but . . . I have always been the open one. (Off-Campus Group)
\end{quote}

This conversation highlights that innocuous behaviors such as discussing masturbation or being “open” (though not sexually active) were sufficient to characterize the young women as ‘sluts.’ This makes little sense if the label is truly applied as a result of behavior – neither of these behaviors are “being promiscuous.” But if, instead, the label is applied on the basis of their gender it all falls into place. ‘Slut’ constitutes women as sexualized beings; it draws
on and helps to recreate systems of patriarchal power in which women are ranked, labeled, and objectified. Once a woman is labeled a ‘slut’ on the basis of her identity as a woman, her behaviors are contextualized as sexual acts. Thus merely discussing masturbation is viewed as evidence of promiscuity, which serves to confirm that labeling her a ‘slut’ was warranted. (Note, though, the disconnect: this kind of thinking is used to justify both that the label was applied as an insult on the basis of her chosen behavior and that her ‘sluthood’ is a natural characteristic.) This kind of covert misuse of derogatives is especially likely with acceptable derogatives, since a taboo more readily marks a term as a derogative.

III. Derogatives terms and derogative speech acts

As Austin and others have pointed out, we do more with speech than simply put forward information. Speech does things. Speech is an action, and through this action we make real changes to the world. We can command, invite, entreat, persuade, call to order, or any number of things via speech. We can use all sorts of mediums to perform speech acts – sometimes we use gestures, or wear particular clothes. Often we use language. However, the speech act is not simply the words that are issued. Rather, words can be the medium via which we perform speech acts. And as such, they can help shape the performative force of the act. Here, I am going to explore the connections between slurs and speech acts. Most of the work on slurs has focused on how they’re deployed in derogatives: in speech acts that derogate on the basis of group membership. Yet not all derogatives deploy slurs, and slurs aren’t only issued in derogatives.
I’ll again review general features of speech acts. Then I’ll give an account of derogatives and the connection between slurs and this kind of speech act. Next, I’ll sketch out how they function first-, second-, and third- personally. Then, I’ll offer a key distinction between direct and indirect derogatives. I’ll argue that slurs, especially acceptable slurs, are often deployed in indirect derogatives that function as multiple speech acts at once. I’ll especially focus on insults and descriptions as forms of indirect derogatives.

Each speech act has a particular performative structure. Not only are speech acts distinct from one another based on who can issue them and to whom they can be directed, but they differ by their function and what they do. A command, issued with the proper authority, in the right context, and to the right person, creates a requirement for that person to do something. On the other hand, an invitation creates an option for the audience to take up or not, and both are appropriate responses. An invitation to coffee between friends doesn’t automatically require attendance – the whole point of an invitation is that it is optional, and we’d be dismayed to find out a friend had come because they felt they had to. An invitation isn’t simply a weak command or one issued by someone with only marginal authority. Invitations and commands have different performative structures and pragmatic outputs. Importantly, we can perform multiple speech acts at the same time: “Gee, I sure am thirsty,” can be an assertion of the speaker’s thirst, as well as a pointed request for someone to get them some water. The surface grammar of an utterance doesn’t necessarily determine what speech act is being performed. When a teacher tells their students, “Please turn in the paper at noon on Friday,” they are not, in fact, issuing a request but rather are commanding their students to get their papers in. We can’t just look to the words being said to determine what action is being performed. We also have to look at the standing of the speaker, the audience,
the context in which it’s being performed, and relevant norms that structure the pragmatics of what’s happening. Importantly, it’s not just about the speaker’s intentions that determine the force of their action. We rely on complex sets of discursive conventions to govern how we take up speech acts (Kukla 2012).

Derogative speech acts target and reinforce social groups, make salient and essentialize particular perceived characteristics of that group, signal and authorize others to take up dehumanizing or derogating views of members of the group, draw on a history or established system of oppression of members of the group, and constitute particular normative implications towards members of the group. Derogatives paradigmatically deploy slurs as part of their surface grammar. This has facilitated blurring the two together.

Derogatives need not utilize slurs. “You people,” in the right context, can work as a derogative without having any words that would typically be classified as a slur. Similarly, spelling out in excruciating detail the perceived dehumanizing and essentialized characteristics of the group, and the subsequent “appropriate” treatment of group members, is a derogative. However, utilizing slurs does two important things. First, it serves as a cue to the audience that the appropriate uptake of the speech is as a derogative. Slurs can facilitate the appropriate uptake of a derogative, and can do so particularly adeptly and strongly. Secondly, slurs are an especially powerful and efficient way of bringing about the intended pragmatic effects of the derogative. Slurs can elicit the dehumanizing ideology and perspective covertly, without explicitly laying out the entire view. Because so much is packed into a single word rather than spelled out, slurs enable derogatives to pack a lot into a small utterance, which makes it that much harder for someone to note, unpack, and resist all that’s been done in the speech act.
Derogative speech acts can be structured first-, second-, or third-personally, and each does importantly different work. First-personal derogative speech acts are those that are self-applied by and towards members of the targeted group (“I’m awful; I am such a bitch.”) If a slur is deployed in the derogative, that slur is not being reclaimed. First-personal derogatives are a form of taking on the essentializing and dehumanizing associations that the term is drawing on. First-personal derogative speech acts are a mechanism of internalized oppression, and they authorize (albeit mistakenly) oppressive treatment of the group. Second-personal derogative speech acts (“You are a bitch”), derogatives deployed by outsiders towards members of the group, are at their core a form of interpellation. Interpellation is “the process by which people are constituted as persons with particular locations in social normative space through vocative calls and the acknowledgements they demand” (Kukla and Lance 2009, 181). By directing a derogative at a person, the speaker calls on them to recognize that they are the subject to whom the label is directed and they are called on to take up that label, with its attendant normative implications. Derogative speech acts serve to mark the targeted persons as essentialized and derogated group members, and they call for a recognition from their targets of this status (Kukla 2016). More than this, they work to elicit a subordinate response from their target. Successfully resisting second-personal derogative speech acts is difficult, since the targeted person’s actions are filtered through the essentialized characteristics called to salience by the speech act. Staying silent works as acceptance or affirmation, and speaking against it is likely to be read as “angry,” “hormonal,” “out of control,” or whatever plays into the stereotypes about the group. Whatever the person does is likely to be read as confirming that the derogative really does apply. While any individual derogative has minimal interpellative force, the repetition of
these speech acts can constitute a person as the kind of subject marked out by the derogative term.

Finally, third-personal derogatives are speech acts deployed from one untargeted person to another (“She is a bitch.”) Here, the derogative works to set up how these outgroup members ought to behave towards and interact with those picked out by the derogative. It might set into play norms under which the targeted person is interpellated, without their prior knowledge, as, for example, ‘crazy’ and so epistemically untrustworthy, or as a ‘slut’ and so essentially sexualized and from whom consent to sex isn’t necessary.

Whether they’re structured first-, second-, or third-personally, these kinds of speech acts directly pick out a member or members of the targeted group. These I term direct derogatives. Statements such as “I’m a bitch,” “You’re a bitch,” or “She’s a bitch” all directly identify a member of the group ‘woman’ and characterize her as strident and unpleasant. Indirect derogatives, on the other hand, refer to someone who isn’t presumed to be a member of the targeted group: “That test was a bitch.” These function via metaphor. This works via a two-step process: first, it sets up members of the group as essentialized and bad, just as direct derogatives do. Then, it sets up a comparison between the target and that kind of badness or essential characteristic. Indirect derogatives imply, “this is so bad it is like them.” When someone says, “That’s so gay!” they’re not drawing on a wholly alternate meaning of ‘gay.’ Rather, this comment only makes sense if being gay is stigmatized. It is a way to condemn or ridicule by drawing an analogy between the object in question and the stigmatized group. Indirect derogatives permeate everyday speech, and especially show up in both insults and descriptions. Again, while indirect derogatives need not employ slurs, the presence of a slur serves as a powerful signal for the appropriate uptake of the speech act. Utilizing a slur, here
too, is both efficient and enables the work of the indirect derogatives to slip by with little notice.

An insult is a speech act that scorns or mocks the person as a result of their behavior. An under-discussed phenomenon in the slurs literature is the way that slurs are deployed in insults towards people who aren’t thought to be members of the derogated group. White video game players, for example, often use racial slurs against one another. These gamers aren’t making a mistake about the racial identity of their opponents. Rather, they are using the slur in order to insult one another by drawing a comparison between the target of the insult and the group identified by the slur. That is, this kind of speech acts has the dual force of both an insult and an indirect derogative. Just because the term is being used as an insult doesn’t mean that the term, itself, stops being a slur. Rather, it continues to convey contempt for the group associated with the slur, to essentialize that group, to license a certain ideology and perspective, and to set up and maintain normative relations with members of the group. It’s just that all of this happens to someone other than who the slur is directed at. Indirect derogatives only work because the term that is being used remains a slur. This is what powers the insult -- otherwise, the insult wouldn’t properly be able to insult.

Because the slur is used against someone who isn’t perceived to be a member of the targeted group, these kinds of speech acts are often excused as ‘just’ insults. They’re considered relatively innocuous. Yet these speech acts are still doing the harmful work of the derogative. They’re simply doing it in a more roundabout way.

Even more innocuous seeming are mere descriptions. Descriptions are familiar speech acts; they make assertions about features of the world and typically aren’t normatively
weighty. Yet when slurs are deployed in descriptions, they often work as indirect derogatives. Slurs targeting mental illness and neuro-atypicality are often utilized in indirect derogatives. Statements like, “Ugh, that test was crazy” are common and go almost entirely unremarked. Yet there’s good reason to think these statements work as more than mere descriptions. A 2007 study found that using mental illness terms as general negatives (as insults or as a way to say, “that’s bad”) made middle school students less likely to seek out mental health resources (Rose, et al. 2007). Even though the mental illness epithets weren’t intentionally or knowingly being directed towards people in the targeted group – and may not have been directed towards people at all – still the students were reticent to take any steps that might indicate they were part of the group marked by the term. That is, there’s some reason to believe that using mental illness epithets in indirect derogatives increases stigma of the group associated with the term in a way that is familiar with how direct derogatives function. The key difference, again, is that indirect derogatives insult or describe something by denigrating a (perceived) separate group.

Indirect derogatives tend to be more socially sanctioned uses of slurs. Because they work covertly they often go unchecked or unrebuked. Because the work they do is masked, they tend to occur more frequently and in contexts where direct derogatives wouldn’t be acceptable. Because it’s not read as a factual description of the target, indirect derogatives are treated as more frivolous, as less serious. So they’re used commonly, and in contexts where a direct derogative wouldn’t go unchecked.

This sets up a feedback cycle: While the direct derogative “powers” the indirect uses, these indirect uses are what keep the term “alive” – if we stopped using the slur as an insult or a description of something that’s extreme or bad, then we would also stop so
continuously reinforcing that this is a ‘bad’ thing to be. Put another way: the everyday deployment of slurs is part of what makes them so powerful and sets it up as being ‘bad’ to be a member of the targeted group, but at the same time the mechanisms by which this everyday deployment takes place masks that this work is going on.

IV. Changing words, changing worlds

The meaning and pragmatics of words change through time. Sometimes this is a natural phenomenon and sometimes it’s intentionally driven. Especially in the case of slurs, the reasons and context for these changes is revealing. There are at least three ways that slurs change: linguistic shift, intentional replacement of the term, or reclamation projects.

Sometimes slurs simply drop out of usage or become neutral descriptors. There are at least two reasons why this might happen: first, it may be because the stigma against the associated group dissipated or diminished. ‘Boche,’ a slur for Germans, was popular when the US was at war with Germany, but fell into disuse once stereotypes and stigma about Germans went away. Second, a slur may fall into disuse because the taboo against the term successfully dissuaded speakers from utilizing it.

Once a term has been identified as a slur that term typically becomes taboo. If it was a technical or medical term, a new, hopefully neutral, term replaces it. However, in what Pinker termed the “euphemism treadmill,” given time that new term often also becomes a slur for the same group. Derogative speech acts, as we’ve seen, needn’t include a slur or derogative term. But if a particular term is deployed consistently in a derogative speech act,
it can come to be associated with the perspective or ideology conveyed by the speech act. That is, using a neutral term in a derogative can, over time, constitute it as a slur.

Some people point to the euphemism treadmill as reason to stop focusing on dismantling slurs. Their argument is that a new slur will always or perhaps naturally crop up, and so creating and enforcing taboos against slurs is futile. However, this is neither inevitable nor unavoidable. If people stop issuing derogatives about the group, the “euphemism treadmill” will likely be interrupted. More than that, instituting taboos against the original derogative serves productive purposes: it disrupts the covert work of the acceptable derogative, and shifts it into a taboo status. This means that the work of the derogative is less likely to slip by people without notice. Second, taking away the original derogative means that those who don’t intend to derogate the group are less likely to do so unintentionally. It usually takes specific and intentional work to issue a derogative speech act when it does not include a derogative term. Instituting a taboo against such terms means that well-intentioned but people are less likely to use such terms unwittingly – for example, even if a benevolent speaker doesn’t understand why ‘colored’ is a derogative, the taboo can help dissuade them from using it. And derogative terms that rely on context are less likely to be unintentionally utilized in a successful derogative speech act – or at least, not at first. An unwitting speaker is simply unlikely to use a contextual term like ‘basic,’ for example, in a way that can be read as a derogative speech act.

Perhaps the best-known way derogatives shift is through reclamation projects. Reclamation projects are intentional efforts to change either the meaning or valence or a term. However, what counts as a “successful” instance of reclamation is vague. One possibility is that the bar for successful reclamation is that the term can’t be deployed as a slur
in an intelligible way. Here, it would be the case that there aren’t the erstwhile sets of pernicious implicit biases for the derogative to elicit anymore. While it still could, technically, be deployed in a derogative speech act, the derogative term would function the same way as any other descriptor and the speaker would have to do a lot of work to get uptake as issuing a derogative rather than an insult. Fully removing the possibility for the derogative term, itself, to issue derogation is a high, and perhaps impossible bar to set for reclamation projects.

A more achievable goal would be for the derogative term to become highly discursively contextual, such that it takes obvious derogatory intent from the speaker to receive uptake as issuing a derogative. This matches well with the trajectory of ‘queer.’ ‘Queer’ is the best known and most successfully reclaimed term to date. It’s used widely in academic and LGBTQIA settings as a pure descriptor. Yet it still can be intelligibly deployed as a slur. Some have pointed to this as a failure of the reclamation project. However, the expanded conceptual taxonomy of derogatives helps us better pinpoint what’s going on here: the reclamation project shifted the term from having derogatory autonomy to being highly discursively contextual. Those wishing to deploy the term in a derogative speech act now must make their intent clear, whereas before they needn’t have. This means that those issuing derogatives no longer have such a ready to hand mechanism by which to signal the appropriate uptake of their speech.

Drawing on the expanded taxonomy of derogatives is helpful for thinking about just what, exactly, a given reclamation project is doing. Is it attempting to alter the essentializing associations connected to the slur? If so, which ones? Or is the project aimed at altering the derogation attached to the slur so that it’s now a term of approbation? The SlutWalk
reclamation project, which I will discuss in more detail in the following chapter, has long baffled those who work on slurs: while self-labeled as a reclamation project, the movement neither offered an alternate meaning for the term ‘slut,’ nor did it give any indication that it was attempting to rephrase it as a positive term. Rather, the SlutWalks are directed at resisting sexual violence and victim-blaming. Now, this reclamation project is intelligible: it is focused on shifting the normative implications of the term so that those labeled ‘sluts’ are no longer read and treated as the sort of person from whom consent to sex is unnecessary. Intervening on these features disrupts the ease with which the slur can be deployed in a derogative. This can help to break the cycle by which slurs both draw on and authorize essentializing, dehumanizing, and oppressive social structures.

V. Conclusion

Derogative terms are a richly varied category. Marking out the conceptual terrain of these variations better enables us to understand how these terms function in our social practices. Drawing out these multilayered distinctions helps us to more precisely understand what a particular term is, and is not, doing. This allows us to note the important differences between, for example, derogatives that refer to the same group of people (e.g., ‘thug’ and the n-word) and how these variations, such as the level of acceptability, allows for different sorts of uses of the word. Recognizing the differences between direct and indirect derogative speech acts positions us to make sense of how these speech acts can reinforce subordination of the targeted group from different angles, and in ways that mutually reinforce each other.
And all this puts us in a better position to resist the oppressive power of derogatives in more precise and directed ways, and to make sense of and support others’ projects of resistance.
CHAPTER 4

THE PRAGMATICS OF RECLAMATION PROJECTS

Derogatory terms, or slurs, are emblematic of broad social practices of oppression and injustice, and at the same time, these terms can themselves be powerful mechanisms of subordination (Camp 2013; Croom 2011; Himma 2002; Hom 2008). There has been much philosophical debate about how these terms work, but there is general agreement that they both draw on and re-entrench deeply embedded stereotypes about the targeted group. Re-appropriating slurs has long been a strategy of groups seeking to fight back against these systems of social injustice. Reclamation projects seek to detach the derogation from the term so that the word may be used as either a bare description of the targeted group or as a term of approbation.

Contrary to what the name of the project seems to imply, reclamation projects aren’t an attempt to ‘take back’ a term the targeted group once had control over; it is rarely the case that the group once had this kind of linguistic control of the term. Instead, reclamation projects are centered on the implicit idea that the targeted group ought to have control over the term that has been used against them. Reclamation projects are a form of social protest, one which is explicitly discursive in nature; whereas other kinds of protest use language as a tool in speeches, songs, or literature to achieve their goal, reclamation projects are focused on changing the linguistic role of a term or phrase. This relies on changing the discursive conventions connected to the term so that the audience can appropriately take up the speech act in which the term is deployed.
Reclamation has frequently been a tool of protest. Perhaps the most well known example today is the largely successful re-appropriation of the term ‘queer’ (Brontsema 2004). From the late 19th century onwards, ‘queer’ was used as a slur for those presumed to deviate from the norms of heterosexual behavior. Beginning around 1990, US lesbian and gay communities began an intentional effort to reclaim the term. Today, ‘queer’ is broadly recognized as referring to an individual whose sexual or gender identity doesn’t fit into traditional binary categories, or someone who rejects heteronormative identity categories. This use of the term has become widespread in general culture, and has made its way into academia with fields such as Queer Studies and Queer Theory. While ‘queer’ is recognized as one of the most successful cases of re-appropriation, it is by no means the only case of attempted reclamation. There have been attempts to reclaim terms targeting ability status, weight, sexual identity, political affiliation, nationality, ethnicity, and race.

Recently, SlutWalks around the world have focused on reclaiming the term ‘slut’ as a method of deconstructing gender norms that excuse or normalize sexual violence. SlutWalks are a protest movement occurring in cities around the world. The movement is centered both on formally organized rallies protesting sexual violence as well as on the dispersed reclamation of the word ‘slut.’ SlutWalks began in Toronto as a response to Police Constable Michael Sanguinetti, speaking at a university health and safety event, saying, “women should avoid dressing like sluts in order not to be victimized.” On March 3, 2011 over 3,000 people turned out for the first SlutWalk to protest the attitudes reflected in Constable Sanguinetti’s remarks; since then, SlutWalks have been held annually in cities all around the globe (SlutWalk Toronto 2011).
The intent behind the SlutWalks, and reclamation projects more broadly, is complex. As with all protest movements, the organizers need to navigate conflicting concerns. First, the express purpose of protest movements is to challenge the status quo. However, in doing so the organizers have to be careful not to fully alienate outsiders or the movement is unlikely to achieve its ends. Second, protest movements need to be about something in order for people to rally around a cause. Yet in doing this the movement risks creating an essentialist construction of the group identity. Third, and perhaps most importantly, the movement needs to direct its limited energy and resources towards the goals that are both achievable and impactful. Yet the more impactful the goal, often the more difficult it is to achieve.

Reclamation projects focus on slurs as a way to navigate these concerns. Reclaiming the slur provides a tangible objective for the group to rally around. Slurs can be an entry point into social practices that license and support embedded systems of oppression. Slurs are emblematic of these broader social systems and use of a slur is itself an instance of subordinate. Reclaiming a particular slur, then, serves as a way to take away a tool that is used to demean and marginalize the targeted group. The slur has historically been used against the group as a whole, and so it’s in everyone’s interests to take away this harmful tool. Reclamation projects focus on changing social norms and discursive conventions: this kind of project can be massively impactful. Yet by focusing on slurs, it avoids overt confrontation with hegemonic power structures that often prompts crushing backlash from outsiders. While reclamation of a derogatory term doesn’t dismantle the entire oppressive system, it may nonetheless be a powerful step along the way. These projects, however, are not without controversy.
Critics, from both outside of and within the targeted group, raise concerns about whether reclamation is the appropriate project on which to focus limited resources of time, energy, and social capital. Some critics worry that reclamation isn’t possible to achieve. These critics fall into two camps: those who believe the project is a misuse of resources because reclamation won’t succeed because the terms are too deeply socially entrenched to be able to alter their meaning or connotations, and those who believe reclamation of the particular slur is too practically difficult to be attainable. If the first position is correct, then reclamation projects may succeed in things such as increasing group cohesion and raising awareness of oppressive social norms targeted at the group, but will never succeed \textit{qua} reclamation of the term. I believe this position is misguided; the seemingly successful re-appropriation of the word ‘queer’ lends weight to this belief. However, I’m not going to give an argument for the conceptual possibility of reclamation here. Rather, in this chapter I’m going to start from the assumption that reclamation is at least conceptually possible. This assumption, however, still leaves open the concern from the second position, that reclamation of a particular term is too practically difficult as to be unattainable.

Whether reclamation is practically attainable, and beyond that, whether it is the proper subject of a movement’s resources isn’t a simple question. The answer to this will always be deeply contextual and necessarily dependent on the specific slur being reclaimed, those engaging in the reclamation project, and the complex web of social norms, practices, and structures in which the targeted group and the term itself are enmeshed. In order to fully understand the potential benefits and risks of reclamation, we need a more complete understanding of the performative structure of reclamation projects themselves. Only with a
robust understanding of what reclamation projects are doing, and their very real potential for both benefit and harm, can individual reclamation endeavors be properly evaluated.

In this chapter I lay out the performative structure of reclamation projects. I argue that reclamation projects face hazards distinct from more familiar types of protest. Whereas other forms of protest may fail to bring about their intended result, instances of reclamation projects can fail to be understood as transgressive acts at all. Even when not understood as transgressive, though, the speech act still has performative force. The way the act is taken up determines the performative force of the act, even when this force is contrary to the original intent of the speaker. When attempts at reclamation fail, context and convention lead the audience to give uptake to the speech act as deploying a traditional use of the slur. The force of this traditional use is to validate and re-entrench the very norms the act was intended to subvert. This is the precarious structure of reclamation projects: when successful, reclamation is the subversion of powerful mechanisms of oppression, but when unsuccessful, the act has the ironic force of constituting mechanisms of oppression. Reclamation, on my account, is conceptually possible, practically difficult, and pragmatically precarious.

I. Socially Embedded Slurs

At their core, slurs target individuals or groups of people in virtue of their membership in a social group. As Iris Young notes, membership in the social group helps to constitute the group members’ identities; it is in being treated as a group member that one comes to be a member of that group, and this shared life experience of being treated as having a particular
social identity helps to bind the group together (Young 1990). As I discussed in the previous chapter, there can be a great deal of variation in whether these social groups are thickly or thinly constituted and how socially entrenched they are. Social groups can range from the highly contextual – membership in a club or sports team – to the well recognized groups organized on the basis of race, gender, sexuality, gender identity, class, age or other identity categories. That slurs target people *qua* social group membership is one of the features that differentiate slurs from insults.

It’s well recognized that slurs do something more than simply refer to members of the particular social group, though exactly what that ‘something more’ may be is contested. Tirrell has argued that slurs license certain derogatory inferences about a group, while Camp proposes that these terms signal allegiance to, and license others to take up, a particularly hateful perspective (Camp 2013; Tirrell 1999; Tirrell 2012). Potts and Williamson, among others, hold that the nature of slurs is best understood as a matter of conventional implicature, and Anderson and Lepore together argue that this ‘something more’ should be understood in terms of slurs’ taboo status (Potts 2005; Williamson 2009; Anderson and Lepore 2013). I offer my own view in the preceding chapter. Here, I am interested in the ways in which these terms fit in to our social practices. Regardless of the particulars of the view, there’s general agreement that slurs are importantly connected to social norms that play a central role in the subordination of the targeted group. That is, slurs both draw on and re-entrench oppressive social norms.

Slurs draw on social networks of oppression, discrimination, or subordination. Tirrell argues that this social embeddedness serves as one way to differentiate deeply derogatory terms from other negative terms (Tirrell 2012). One way to illustrate this is through the
contrast between the N-word and ‘whitey’ both are pejorative terms targeting individuals in virtue of their membership in a racial group. Though ‘whitey’ is certainly pejorative, it doesn’t connect up with a long standing practice of systematic dehumanization, marginalization, and exclusion from participation in social spaces. The N-word on the other hand, does have such a history and its use continues to draw on and endorse such practices today (Jeshion 2013; Rahman 2012). On this account, ‘whitey’ fails to be a deeply derogatory term, while the N-word is one.

Exercitives, as discussed in Chapter 2, “Identity and Discursive Signaling,” are a type of speech act that make rules binding. They set permissibility conditions for certain sorts of actions in particular contexts, especially insofar as they confer or deny rights to a person or group of people. While exercitives were initially conceived of as setting these permissibility conditions (granting voting rights or membership in a club, for example), Rae Langton and Mary Kate McGowan both have written on the unofficial yet still powerful way speech can create norms that allow for or block off possible moves in social space for particular people or groups (Langton 1993; McGowan 2012). Slurs deployed in speech acts often function as exercitives; they create and re-entrench social norms under which members of the targeted group are subordinated, marginalized, or degraded (McGowan 2012). So, Police Constable Michael Sanguinetti’s remark that, “women should avoid dressing like sluts in order not to be victimized,” both drew on systems of patriarchal power under which women are subordinated to men and are denied sexual agency, and it helped to constitute this social

---

10 I’m using the practice of mentioning the slur by its first initial as a way to signal that I do not think even mentioning the slur is acceptable. As I discuss in Chapter 3, “Talking About Slurs,” even mentioning slurs in this oblique way may well constitute harm to readers. I hope that the value of this example justifies that harm here.
reality in which women are subordinate to men, women who dress or behave in so-called ‘provocative’ ways are denigrated, and survivors of sexual violence are cast as responsible for their own victimization.

Of course, an individual utterance typically has limited normative power. While some deployments of slurs are issued by speakers inhabiting official positions of authority, and thus are able to wield far reaching power, most often slurs are deployed in everyday non-official speech acts (see Maitra 2012 for a discussion of the role of authority in subordinating speech). The real power comes from the ubiquity of awareness and usage of the slur. Butler has pointed out that it’s the repetition of an act that constitutes it as a norm and as normal. This persistent use constitutes the slur as having entrenched social power (Butler 1997). The more the slur is deployed, the more power it has and the greater power is has to draw on, thus further re-entrenching the slur’s powerful status.  

Slurs also license further harmful behavior. Lynne Tirrell describes this feature of derogatory terms as being “action engendering” (Tirrell 2012). The social norms constituted by use of the slur are the scaffolding within which we perform our daily activities. These norms act as a lens through which we see the world, process information, and determine appropriate and desirable behaviors. Use of slurs allows for and endorses certain behaviors that might not otherwise be salient options. Sanguinetti’s remarks license particular kinds of behavior towards women dressed in so-called ‘revealing’ attire. Rather than preventing sexual violence, Sanguinetti’s comment casts the victims of sexual violence as responsible for the

---

11 This can be temporally extended: so long as we, as a society, are broadly aware of how a slur has been historically deployed it retains its strength, while if we forget about the slur’s history that strength dissipates.
violence inflicted on them based on having ‘dressed like a slut.’ This kind of victim blaming helps constitute women who are perceived to be ‘sluts’ as the appropriate targets of sexual advances and opens up real possibilities for sexual violence to be enacted upon these women. This pervasive threat of violence, whether or not it is instantiated in any particular instance, is itself a form of oppression (Young 1990). When a slur is deployed, it reifies this threat of violence.

So, slurs deployed in speech acts draw on entrenched systems of social power, constitute oppressive social norms, and open up real possibilities for harmful action directed against those picked out by the slur. Reclamation of the slur serves as a way to disrupt these entwined avenues of social power. Reclamation projects take away a tool used for expressing oppressive power against the targeted group, they remove a method of constituting oppressive norms, and they block off one way by which harmful actions are licensed against the targeted group. As the reclaimed use of the term becomes more widespread, it further weakens the connection between the term and the pernicious norms that governed its force. Reclamation, then, is the subversion of oppressive social norms.

Reclamation, however, can’t happen in a vacuum. These projects can only occur in conjunction with other forms of protest and subversion. Various forms of social change work together in a feedback loop: changing the norms helps to change the discursive role of the term; changing the discursive role of the term helps to change the norms. This is well evidenced in the reclamation of ‘queer:’ US lesbian and gay activists engaged in various strategies to change the prevailing view of homosexuality. This allowed space for ‘queer’ as a positive term. ‘Queer’ as an identity descriptor then helped to normalize non-heteronormative identities.
Slurs, then, are deeply and intricately socially embedded. They draw on, constitute, and support oppressive norms, and reclaiming a slur is the subversion – at least in part – of these norms. However, though this is important groundwork, it doesn’t yet shed light on the performative structure of reclamation. In order to see this performative structure, we first need a better sense of the performative force of speech acts.

II. Performative force and discursive distortions

The performative force of a speech act is the normative change wrought by that speech act. As noted earlier, speech acts are multifaceted: a single speech act can do many things. The performative force of the act encompasses all the normative change wrought by the speech act, both the changes occurring in the moment of the utterance and those that occur farther down the line. Remember, Austin argues that speech acts have three parts: the locution, or the act of uttering the words, the perlocution, or act of bringing about effects by the utterance, and the illocution, which is the type of action constituted by the utterance (Austin 1962). In speech acts deploying slurs, the slur is a part of the locution of the act. The locution of the speech act sometimes, but not always, points to the type of illocutionary act performed by the speech act. “Please turn your papers in by Friday,” has the surface grammar of a request, but when issued by a professor to her students is properly a command: the professor has the appropriate authority for this illocution, and so her students have only one normatively appropriate option: to adhere to the perlocutionary force and turn in their papers by the due date. If, on the other hand, the professor says to her friend, “Please come to my house for dinner tonight,” though the surface grammar looks similar to the order to the students, in
this case she has issued an invitation, one that her friend can appropriately accept or decline. As seen here, the type of speech act doesn’t always determine the perlocutionary effect of the speech act; a given illocution can lead to different perlocutionary effects. Similarly, various kinds of illocutions can be used to achieve a single desired perlocutionary end: the professor can order her friend, or entreat her friend, or bribe her friend, all to achieve the desired outcome of having that friend at dinner. Finally, the force of speech acts can be conjunctive. Sanguinetti’s illocution was both an exercitive and a declarative, and the perlocutionary effects are still reverberating in the form of SlutWalks.

In many ways, the force of a speech act is outside the control of the speaker. Austin argues that certain felicity conditions must be in place for an utterance to achieve its intended purpose. To borrow an example from Austin, when a duly designated individual at a ship christening ceremony says, “I name this ship the *Queen Elizabeth,*” that ship is now named the *Queen Elizabeth,* while it wasn’t so before (Austin 1979). However, if “some low type” were to run up to the ship, grab a champagne bottle out of the duly appointed person’s hand, smash it against the ship and say, “I name this ship the *Generalissimo Stalin,*” it would not be the case that the ship is called the *Generalissimo Stalin.* The interloper did not meet the necessary conditions for christening the ship, and the speech act did not have the intended performative force of christening the ship with a new name. So, certain conditions must be in place for a speaker to achieve the intended force with their speech.

While Austin’s account emphasizes the role of intentions in determining the force of a speech act (assuming the felicity conditions are met), I follow Langton, Rebecca Kukla, and Mark Lance, amongst others, in holding that intentions play less of a determinative role than Austin thought. While certain conventions must be in place for an utterance to be properly
taken up as a particular type of speech act, it’s this uptake that determines the performative force of the speech act. Both the illocutionary and the perlocutionary force are dependent on how others take up the utterance. That is, a speech act can only do things insofar as people recognize it as doing something and act on that recognition. Of course, though, uptake of a speech act isn’t simply up to the whim of the audience. Rather, we rely on a complex set of conventions, such as whether or not the speaker has the appropriate standing to issue the act, to give meaning to linguistic moves (Kukla 2012). Social context and discursive convention play significant roles in how a speech act is taken up. These conventions govern when something ought properly be taken as a command or request, as an assertion or an expression of emotion. Awareness of micro contexts further helps us to refine our discursive sensibilities: subtle norms brought to salience in a particular context may shift the proper uptake of a speech act.

Just because an utterance is properly taken as a particular speech act doesn’t always mean it will be. Sometimes the audience makes an innocent error – perhaps because they misheard what was said or because they missed the subtle cues conveyed through tone or body language. Other times, however, speech isn’t taken up as the intended action due to what Langton terms *illocutionary disablement* (Langton 1993). Here, a speaker is unable to receive uptake as performing the intended type of speech act because of norms that block off her performance of that action. The speaker meets the official felicity conditions for issuing the act, and properly ought be taken as performing it, but the speaker is pragmatically unable to

---

12 While I use Langton’s original terminology in this paper for purposes of clarity, in general I recommend referring to this concept as ‘illocutionary dis-enablement’ so as to avoid ableist connotations.
do so because of social norms that alter the audience’s uptake. Langton uses pornography as her illustration of this phenomenon: subordinative pornography, she claims, constitutes social norms by which women’s sexual refusals cannot be recognized as such. On this account, pornography is a kind of speech act and the illocutionary force of pornography is to constitute social norms under which women lack sexual agency. These norms set discursive conventions whereby women’s sexual refusals are unintelligible as refusals. It’s not simply that the refusals are unheard or ignored, but rather that they aren’t recognized as refusals at all. When this occurs, Langton argues, women have been blocked off from issuing refusals – they have been silenced.

Langton is quite right to say that in these cases a speaker is unable to do what she intends with her words. However, she is blurring together two importantly distinct kinds of illocutionary disablement: illocutionary silencing and illocutionary distortion. For the most part Langton, and the many others who have taken up her work, implicitly focus on illocutionary silencing. In these cases the speech act isn’t taken as any kind of illocutionary act at all: it has no performative force. In the case of sexual refusals, the speech act is neither unheard nor ignored, yet it doesn’t receive uptake from the audience as anything – it’s akin to screaming into the wind.

Yet not all failed speech acts completely lack performative force. Often, speech still does things – including powerful and important things – even when the changes wrought by the act are not those intended by the speaker. These are instances of what I term illocutionary

\[13\] For further discussions of Langton’s arguments regarding pornography and silencing, see: Horsey and Langton 1998; Langton and West 1999; Bird 2002; Maitra 2004; Bianchi 2008; Maitra 2009; Dotson 2011; McGowan, Adelman, et al. 2011; Mikkola 2011.
distortion: when a speech act is taken up as a different type of illocutionary act than that intended by the speaker. To return to the case of sexual refusals, pornography can set up gendered norms in which women are understood to ‘play hard to get’ as a form of sexual encouragement. In these contexts, a woman’s sexual refusal isn’t taken as nothing, but instead as a kind of encouragement and participation. That is, the force of her speech is coopted and turned against her, and to claim she is silenced paints an incomplete picture; it ignores the real and harmful performative force of her words. To have her speech turned against her is an importantly different kind of disempowerment than to not be recognized as issuing a speech act at all. We need recognition of both illocutionary silencing and illocutionary distortion in order to have an appropriately rich conceptual framework of the ways in which illocution can go astray.

Most of the time the speaker’s intent aligns with the uptake the speech act receives. Sometimes, however, the audience must navigate competing discursive norms: different sets of norms yield different understandings of the illocution performed (Kukla 2012). So, to return once again to Langton’s example, explicit norms about the right of sexual refusal yield one kind of uptake of a woman’s intended refusal, while the social norms constituted by subordinative pornography yield an entirely different result. Sometimes features of a speaker’s identity will do this work of calling to salience one set of discursive norms over another (Kukla 2012). Here, the woman’s identity as a woman intersects with the norms constituted by the pornography to lead the audience to take up her refusal as encouragement. The point is that both official and unofficial social norms and practices work to govern our discursive conventions, and these various sets of practices can come into conflict. In these cases where the governing discursive conventions are underdetermined,
distortions are especially likely. Reclamation projects necessarily take place in these contexts of underdetermined discursive norms.

III. The performative structure of reclamation

Reclamation projects are made up of myriad individual speech acts in which the slur is deployed; in these acts the speaker intends for the slur to be either a neutral or positive descriptor. It’s not enough for the speech act to be intended as a positive; what matters is the way the term is being deployed in the speech act. So, while “Don’t worry, you’re not a slut,” is a positive reassurance, this deployment of ‘slut’ relies on the traditional discursive role of the term. In contrast, while “It’s really gross that you claimed to be queer when you’re not just to try to fit in at that activist meeting” expresses a negative sentiment, ‘queer’ is being deployed as a bare descriptor of non-heteronormative identity. What matters is the intended role of the term in the speech act. Sometimes these acts are overtly positioned as part of a reclamation project and sometimes they are not; speakers don’t need to explicitly intend to participate in reclamation projects for their speech to be part of the movement.

These projects are directed at altering the discursive conventions governing the audience’s uptake of speech acts that deploy slurs; this is the far-reaching perlocutionary effect of reclamation. Once a reclamation project has fully succeeded, speech acts deploying the erstwhile slur won’t have oppressive force in virtue of deploying that term. They will cease to be exercitives constituting oppressive norms, and mere deployment of the slur won’t license harmful behavior directed at members of the social group. The purpose of
reclamation projects, as such, is to bring into being new discursive conventions governing uptake of speech acts deploying the slur.

New discursive conventions don’t simply spring into widespread usage; they need to be brought into being and propagated. The illocutionary act of reclamation is to constitute these new conventions and is an act of transgression against the existing social norms. In order to bring about these new discursive conventions, speakers must deploy the slur in contexts where the new conventions have not yet taken root. Typically, this means deploying the slur either in conversations with people outside the group seeking to reclaim the term or in contexts where these outsiders may encounter, and give uptake to, the speech act. In these contexts, the audience will not yet have the non-oppressive discursive norms to govern the uptake they give the speech act. Tirrell points out that in these instances the audience will take the speaker to be using term in the traditional way (Tirrell 1999). While this may be true, it paints an incomplete picture: it’s not merely that the term is being misconstrued, but that the speech act in which the term is deployed is distorted. In these instances, regardless of what else the speech act is doing, part of its performative force is to constitute mechanisms of oppression.

I will argue in the following chapter that the mere presence of an unreclaimed slur, regardless of how it is discursively embedded, can elicit dehumanizing associations about the targeted group. These implicit associations are part of the perlocutionary effects of the speech act. They also play a role in the exercitive force of the speech act in which the slur is deployed – priming and reinforcing these dehumanizing associations about the targeted group is part of how oppressive norms are sustained. Successful reclamation entails changing the associations primed by the slur – either changing the power of the slur to elicit those
effects or changing the sets of associations called forth in the first place. This, in turn, can help depower the erstwhile exercitive by curtailing the normative force of the speech act.

Some audiences will be able to recognize that speakers engaging in reclamation are doing something atypical with their speech. Lacking other guiding conventions, though, these audiences will often continue to rely on already entrenched discursive conventions to structure uptake of the speech act as the oppressive exercitive of old. Other times, speaker identity will help to resolve discursive under determination: if the speaker is a member of the group targeted by the slur, this may help to call into salience the intended set of conventions and alternate set of associations connected to the term so that the audience may give the intended uptake to the speech act. It’s not simply that the speaker’s identity gives them warrant to issue the erstwhile slur, as Anderson and Lepore propose, but rather that due to this identity the speaker may be positioned to perform speech acts that others cannot (Anderson and Lepore 2013). Yet even this isn’t assured, and especially not in the early stages of the project when the nascent set of new conventions is largely unknown. In these cases, despite the intent of those engaging in reclamation, the force of their speech is distorted; rather than subverting the pre-existing social norms, their speech will continue to constitute the oppression of those targeted by the slur.

One might hope to allay this concern by performing reclamation only within the discursive community trying to reclaim the term. In these contexts, most people know of the intended new conventions, even if those conventions have not yet fully solidified. The explicit knowledge of and commitment to the aims of the reclamation project ideally override preexisting discursive norms. In these contexts, the audience would be able to properly give uptake to the speech act as transgressive. On this model, reclamation would
only take place within an isolated discursive community; either speakers would fall back to adhering to the traditional conventions when interacting with outsiders, or they would not interact with outsiders at all.

Unfortunately, this strategy fails. The first reason for this is practical: few, if any, social groups are entirely isolated, and especially not linguistically. Most people belong to multiple complex and overlapping discursive communities, and move between these communities fluidly everyday. A person might employ professional terminology while with co-workers, slang and references to inside jokes while amongst various friend groups or family, and yet other specialized language while with her sports teammates. For most people the boundaries of these micro communities are blurry: colleagues discuss work while outsiders can overhear them at cafes and bars, friends chat at parties surrounded by lesser known acquaintances and invite newcomers into their conversations, and the lexicon from one micro community is unwittingly deployed in others. Our discursive practices often spill over from one sphere of life to another, whether we intend them to or not. When this happens, our speech is nonetheless heard and taken up, and can impact the audience regardless of whether they are the intended recipient of the speech.

Practical considerations aside, a reclamation project would not be able to achieve the desired ends if it were entirely discursively isolated. A central goal of reclamation is to subvert oppressive social norms used against those targeted by the term. This subversion can only occur by interacting with people outside the group engaged in reclamation. If discursive conventions are subverted only within a particular community, then outsiders will continue to draw on, constitute, and re-entrench these oppressive social norms, which have a profound and real effect on those who are targeted by them. To participate in reclamation
centrally involves performing speech acts in contexts where the discursive rules governing the uptake of the speech act have not yet been changed. This means deploying the erstwhile slur in contexts where either the ruling discursive conventions are in flux or where they lead the hearer to give uptake to the speech act as a traditional deployment of the slur.

This, then, is the performative structure of reclamation projects: the project is composed of many individual speech acts which must, at least some of the time, be performed in contexts of discursive flux. When these speech acts are successful, they have the performative force of subverting mechanisms of oppression. Yet when they fail, they ironically constitute the very thing they were intended to subvert. When this happens, speakers from the targeted group have the force of their speech turned against them; what may be intended as an act of liberatory protest may instead work to constitute harm to the speaker. This risk is unavoidable; it is part of the structure of reclamation. Reclamation projects are necessarily precarious.

IV. The power and limits of reclamation

I began this chapter by raising concerns about whether reclamation projects are an appropriate use of time and energy for a protest group. Even starting from the assumption that reclamation is conceptually possible, the answer to this question is by no means clear. If I'm right, then reclamation projects face particular challenges due to their performative structure. These structural hazards are distinct from traditional concerns about the difficulties of implementation: on that view, if done at the right time, in the right place, and by the right people, presumably reclamation could be a safe and stable affair. The hazards
I’ve pointed out, though, aren’t avoidable: they’re built into the performative structure of the project. When they fail, reclamation attempts have the ironic force of constituting mechanisms of oppression.

This raises concerns for marginalized groups fighting against mechanisms of oppression. Reclamation, as I have pointed out, can only be successful insofar as it is connected to other forms of social protest. If these other forms of protest aren’t already in place, then the attempted reclamation project is likely to have the unintended force of reifying the pernicious norms the group is working to fight against. Even when these other forms of protest are in place, though, reclamation ought to be approached carefully. Since the hazards of reclamation can never be fully avoided, this raises real risks for already marginalized groups. In some cases, the hazard of further entrenching their subordination may be too dangerous to take on. Subordinated groups that have social capital on other fronts, such as the upper-middleclass white gay and lesbian communities who drove the reclamation of ‘queer’, may be situated such that they are able to face these risks – while others may not be so positioned. Additionally, the group targeted by a slur is complex, made up of people with intersectional identities and different kinds and degrees of social power. This point was eloquently raised by a group of Black women responding to the SlutWalk movement: “As Black women, we do not have the privilege or the space to call ourselves ‘slut’ without validating the already historically entrenched ideology and recurring messages about what and who the Black woman is. We don’t have the privilege to play on destructive representations burned in our collective minds, on our bodies and souls for generations” (Anonymous 2011). Not only are slurs directed differently at people depending on the
particulars of their identity, but these same particulars of identity make real, lived differences in the unavoidable risks of reclamation.

Luvell Anderson offers a potential resolution to this problem: while he recognizes that reclamation may unintentionally worsen material conditions of those targeted by the term, he believes rejecting reclamation on these grounds wrongly prioritizes certain kinds of harms over others (Anderson forthcoming). According to Anderson, the harm of slurs can be sorted into two broad categories: increasing external constraints, such as voting restrictions or laws tacitly holding victims of sexual violence accountable for their victimization, and internal constraints, which are comprised of the psychological harms and limits to mental development done to the targets of the slur. Anderson argues that reclamation is a form of resistance directed against internal constraints, and that this ought be given priority – even at the risk of increasing external constraints. While I agree that reclamation can be a way to fight against these internalized harms, I believe this view is too simplistic: when attempts at reclamation are taken up in such a way that constitutes and perpetuates harm against the targeted group, the distinction between internal and external constraints is not clear. Deployment of slurs that sustain harm increases both external constraints, by increasing the real possibility of instantiated violence, and internal constraints, by reifying the psychological damage that comes from the pervasive threat of violence and the knowledge that one is generally seen as an appropriate and consequence-free target of that violence.

The performative force of distorted reclamation is complex and multifaceted, and all of this force needs to be taken into account. The point is not that reclamation ought only be attempted by those who already have social power, but rather that this is one tool, among many others, by which to fight against mechanisms of oppression, and like all tools its uses
and limitations need to be properly understood in order for it to be yielded effectively. Whether or not reclamation is a worthwhile project will depend on many things, ranging from what other forms of resistance are already in place, to the embodied social positions of those who would primarily be engaging in the reclamation, to the degree and kinds of risks already faced by those targeted by the term. In assessing all this, there needs to be an awareness that while individuals endorsing the reclamation project may be situated such that the hazards of the project are something they practically can take on, reclamation projects are broad and the performative force of distorted attempts at reclamation impacts other members of the targeted group who may not be so fortunately situated. Although the predominantly white women who started the SlutWalk movement believed themselves to be able to take on the risks of the project, it’s not clear that the leaders considered the embodied situation of others who are targeted by the term. While ideally all will benefit from the subversion of oppressive norms, to disregard the real potential for increased harm and danger is recklessly naïve.

Reclamation is intrinsically hazardous. Yet it can also be worthwhile: even when reclamation projects fail as reclamation, they still may accomplish good. They may increase group cohesion, bring awareness to the cause, or provide impetus for other sorts of resistance. At its best, reclamation makes real differences in the lived situation of all those who have been targeted by the slur. And when successful, reclamation of derogatory terms is the subversion of powerful mechanisms of oppression. While it will always be precarious, it may on occasion be well worth the risks.
CHAPTER 5

TALKING ABOUT SLURS

While it’s generally well recognized that we oughtn’t use slurs, we’re still unsure just how, exactly, to navigate talking about slurs – what to do in instances in which slurs are mentioned but not used. In this chapter, I’m going to explore an under-discussed aspect of slurs: the connection between slurs and implicit associations. To the degree that philosophers have explored the pragmatics of slurs, they have typically focused on the illocutionary force of slurs. While this is certainly important, the perlocutionary effect of slurs has been mostly ignored. In focusing on the connection between slurs and implicit associations I hope to help fill in this gap. Not only does this help to shed light on philosophical puzzles about slurs, it informs the contours of talking about slurs. I’ll begin by laying out the central puzzles about slurs that philosophers of language have been grappling with. Then, I’ll show how one of these issues – this worry about mentioning slurs – comes up in much broader contexts than just within philosophy of language. Next, I’ll turn to research on priming and implicit associations. I’ll suggest that slurs can act as a powerful priming mechanism of pernicious implicit associations about the group targeted by the slur. If this is right, it raises serious concerns for instances where slurs are merely being mentioned. Based on this research, I offer practical guidance for how we ought to navigate talking about slurs. Finally, I show how this understanding of the role slurs play in priming pernicious implicit associations can shed light on philosophical puzzles about slurs.
I. Puzzles about slurs

Before delving into the central puzzles about slurs, first a reminder of what slurs are. Though there’s substantial debate about just what, exactly, constitutes a term as a slur, there is general recognition that slurs have the following characteristics: Slurs identify members of a social group – often on the basis of race, gender, sexuality, gender identity, nationality, religion, or disability. They essentialize that group, by picking out a characteristic or set of characteristics and taking those to be representative of members of the group. Slurs are typically dehumanizing, or disparaging, or derogating of the people they target. And they’re socially embedded, in that the term draws on and calls into salience a history of oppression of the group. Finally, slurs help to establish norms for the treatment of members of that group.

Most people recognize that it is bad to use slurs. Many have written on what, exactly, makes using slurs so bad. Some argue that slurs license hateful perspectives (Camp 2013), others that they essentialize their targets and set up conditions for harmful action against members of the group picked out by the slur (Tirrell 2012). Some have argued that using a slur is bad because it violates social taboos against usage of the term (Anderson and Lepore 2013), or because the slur carries with it an unavoidably hateful conventional implicature

---

14 I explore these facets of slurs in detail in the Chapter 3, “Derogatives: Insults, Slurs, and Speech Acts.”
15 This is the point where it might make sense for me include a long list of slurs to help illustrate what I’m talking about. But I’m not going to do this, and in fact I’m not going to include any slurs in the body of this paper – for reasons that will hopefully become quite clear by the end.
16 In this paper I am setting aside cases in which a slur is being or has been reclaimed, though the conclusions of this paper can help to elucidate the importance and hazards of reclamation.
(Potts 2005; Williamson 2009). There is little consensus about the exact mechanics of slurs, but throughout these various theories there is overwhelming agreement that it is bad to use slurs and in most cases that to do so causes harm to the audience or members of the group targeted by the slur. That is, we agree that it is bad, though not on why it is bad.

In thinking about slurs, there are several puzzles that have caught philosophers’ attention. Each of these puzzles is highly debated and many answers have been proposed for each one. How one answers these puzzles depends on which theory of language one ascribes to.

First, and most centrally, there’s a great deal of debate about just what, exactly, makes slurs bad. Is it that they cause offense? Or harm? Or is it that they dehumanize their targets? How we cash out the bad-making feature of slurs influences how we understand what it is that slurs do. Another way to put this puzzle is: What distinguishes slurs from their counterparts? Typically, for each slur there’s also a non-slurring way to refer to the group: the category label. The category label is the currently socially acceptable way to refer to the group. Both slurs and category labels seem to identify members of the same group, but the slur does something over and above what the category label does. And there are a great many theories about what that might be.

---

17 Standard convention within philosophy is to refer to this term as the ‘neutral counterpart’ to the slur. This, I believe, is a misnomer – it is not obvious that all such terms are neutral (see: the evolution of ‘colored’ or many medicalized terms for sexuality as supposedly neutral counterparts which have since been recognized as perpetuating a particular oppressive ideology). Arguments that presume the term is neutral in order to draw conclusions about what makes the slur derogatory often end up begging the question. Here I adopt the conventions favored by the investigators in using the term ‘category label’ to describe a currently socially acceptable label for the group, without making any assumptions about the neutrality of such a term.
Second, denying the application of a slur doesn’t seem to work as in the standard way. If a speaker says, “Jamie’s not a [slur],” the speaker still seems to be somehow implicated in whatever it is that the slur is doing – even when the speaker is trying to deny the application of the slur. Many philosophers have noted that denying the application of a slur simply denies that the term doesn’t apply to this person, but there may well be others who are properly categorized by the slur. This is different from saying, “Star isn’t a unicorn; Star is a horse.” Here, standard denial seems to work – the utterance simply holds that Star is not a unicorn, but doesn’t imply either way whether unicorns exist. Just why slurs are different is unclear.

Third, speaker identity is recognized to play a role in how slurs function. It seems to make some sort of difference whether the speaker issuing the slur is a member of the group targeted by that slur. There is a general sense that it’s more acceptable, or less harmful, or has different pragmatic impact, for a member of that group to say a slur than it is for someone not affected by the slur. This extends beyond first-personal discussions of how the slur impacts members of the group of which the speaker is a part and holds for nearly all contexts.

Next, slurs sometimes undergo reclamation. As I’ve discussed already, this happens when, say, the people targeted by the slur start using it as a badge of honor or as a pure description. This occurred with the phrase “nasty woman” during the 2016 presidential election, which was embraced by the people it was intended to demean. However, it’s not entirely clear just what is happening in these reclamation projects, what the success conditions are, or who can engage in them.
Finally, the use/mention distinction seems like it might run into some kind of problem with slurs. Philosophers recognize a distinction between using and mentioning a term or phrase: the difference between saying, “I hate basketball,” and saying, “Alex said, 'I hate basketball.'” In the first case, the speaker is committed to the statement, while in the second they are merely reporting on what another person has said. On the surface, it seems like slurs ought to work in this standard way: while using slurs is bad, mentioning them doesn’t reflect on the speaker – they’re merely reporting on an event or artifact. Yet there’s a general worry that mentioning slurs has at least some kind of residue or is fraught in some way.

II. Talking about slurs

Whether and how use/mention breaks down in the case of slurs poses a problem in a broad array of contexts. From research to journalism to educational contexts, slurs are often mentioned. In research venues, there is ongoing debate about how much, and when, it is appropriate to mention slurs. This comes up in research articles and at talks and conferences. Psychologists studying the impact of slurs include the particular slurs that were studied when reporting on their results. Philosophers trying to understand slurs include examples in their work. At a recent philosophy conference on slurs, a woman repeated a slur six times in the course of asking a question. The worry isn’t simply that academics are used to including such examples in their work and changing this practice may be difficult. Rather, it’s typically held to be a virtue in academic work to be specific and to include examples. If use/mention breaks down in the case of slurs, then this erstwhile virtuous practice may itself cause harm.
Journalists come up against this concern as well. Journalists have to figure out what to do when a public figure utters a slur – do they quote it? Or not? When a politician or celebrity uses a slur, there seems to be good reason to report on what specifically was said. Journalists have a professional obligation to produce precise, accurate news. Yet there is often a massive uproar when journalists do this. Even reporting on things as innocuous seeming as sports runs up against this problem when the team name is itself a racial slur – as is the case with Washington, D.C.’s football team (Halbritter 2014; National Congress of American Indians 2013). There have been numerous news programs devoted specifically to discussing what to do about reporting on slurs (Bershad 2012). This topic is identified as a contentious one within ethics of journalism – but the unhelpful advice that is usually given is to follow the conventions of the publication. While it is recognized as contentious and important, there is not much helpful guidance for what those conventions ought to be (Ahtone 2013; Bershad 2012; ONAethics n.d.. For the official practices of many well known publications, see Maynard Institute 2012).

Finally, there is ongoing controversy over the role of slurs in classrooms, even when they are mentioned for purportedly educational purposes. This issue comes up with increasing frequency and spectacle at universities across the US. This is illustrated well by an exchange between an audience member and a 1st Amendment scholar speaking at Brown University (Flaherty 2017). The scholar gave a public lecture defending free speech and in the course of this he mentioned a racial slur. When a student, from the group targeted by the slur, politely requested that he not say it he responded:

“You can’t talk about the words in the class when you’re discussing whether the word should be legal or not? Doesn’t make any sense. Or you read it in a novel that uses the words and you can’t use the words? Sorry. But I do hear you.”
This reply exemplifies much of the debate about whether or not slurs out to be mentioned in educational settings. Once again, the thought is that it is not only necessary but virtuous to mention the slur, to have it ‘out on the table’ to talk about. Yet others point out that including the slur in this way seems to create some kind of harm, especially for those targeted by the slur. Though the scholar grapples with this, ultimately he, like many others, believes that since he’s merely mentioning the slur he is not enacting harm by saying it. That is, many people suppose that the use/mention distinction serves as a shield – the thought is that so long as the slur is within quotation marks or is merely mentioned, there’s no problem with including it.

Increasingly, articles, conferences, and classrooms in which slurs are discussed include disclaimers that such words will be mentioned, though they do not alter their conventions about saying the slur itself. Though there is certainly no unified move against mentioning them, there is more and more debate about what to do about mentioning slurs.

Now, some might argue that the controversy about mentioning slurs is simply due to a mistake. The audience in many of these cases includes people unfamiliar with the use/mention distinction and the pushback against mentioning slurs may be due to the hearers’ confusion. They thought the slur was being used when, in fact, it was merely being mentioned. If this is true, then there is still a real ethical worry here: that those who talk about slurs are unintentionally being taken as using the slur and by so doing are taken to endorse the slur and all it entails. This is not trivial – harm that results from misunderstandings is real harm nonetheless – but it is one that could be resolved relatively easily. A writer or speaker could include a short disclaimer: “in the following I will report
on/mention/quote slurs, but I’m not actually using or endorsing them.” Even if it takes
more complex scaffolding, the thought is that by clearly identifying what they are and are not
doing the speaker could navigate this ethical problem.

This is a useful move, which I will come back to later. Unfortunately, I don’t believe it
can fully solve the worry about mentioning slurs. While surely sometimes the harm of talking
about slurs comes from this kind of misunderstanding, the real harm goes much deeper and
cannot be resolved so easily. In the following I do not claim to give an exhaustive account of
slurs or the harm of talking about slurs. Rather, I explore one important and underdiscussed
aspect of slurs: that they serve as a powerful way to call forth pernicious implicit associations
about the targeted group.

III. Implicit associations

Implicit associations\(^{18}\) are the automatic, non-conscious, unreflective representations that we
hold. Sometimes our implicit associations align with our consciously held beliefs, but often
they do not. I’m not going to advocate for a specific view of implicit associations here –
rather, I’m going to draw on general understandings of how they work in a way that is,
hopefully, compatible with a range of views. Perhaps the most well-known example of

\(^{18}\) ‘Implicit association’ and ‘implicit bias’ are synonyms. While lay people sometimes
interpret ‘implicit bias’ to mean ‘implicit prejudice,’ this is an incorrect reading (though some
implicit biases are prejudices). Implicit associations (or biases) can be either positive or
negative, and needn’t have anything to do with social prejudices. The ‘bias’ refers to a person
being more readily inclined, or biased, towards some associations over others. If someone
has an implicit association between ‘thunder’ and ‘lightening,’ then when they hear the word
‘thunder,’ they can more quickly call forth the word ‘lightening’ than, say, ‘banana.’ When
they are exposed to ‘thunder’ they are biased towards ‘lightening.’
implicit associations are from the IAT that measures the time between exposure to a black or white face and a picture of a weapon or tool; overwhelmingly, Americans are faster at associating black faces with weapons than with tools. Similarly, Americans are faster at association white faces with positive words and black faces with negative words. So, implicit associations work when a stimulus calls forth a set of associations. These associations can be content based, can be negatively or positively valenced, or (as there is good reason to believe) can be a complex combination of content and valence.

We have little direct control over our implicit associations; they arise from repeated exposure to information or ideas. Once an association has been created, we can be primed to call forth the association. This especially matters when the association pertains to social groups: when it relates to things like race, gender, sexuality, gender identity, and all the myriad intersections amongst these. Implicit associations impact our behavior in myriad ways, particularly in moments when we are stressed, under time pressure, multitasking, or other forms of increased cognitive load. We internalize implicit associations about ourselves which then impacts our behavior: filling out demographic information on a standardized test can call forth associations about race and gender and academic performance. In the case of white men, this is often a boon and helps them to do better, while in the case of most women (for math and science at least) and for many black and Latinx people this can lead to decreased academic performance on the test.

Implicit associations can impact how we interact with other people. One of the studies that illustrates this well shows that if a backpack is put in a corner of a room, people are more likely to collaborate, whereas the presence of a briefcase is more likely to lead to competitive behaviors (Kay, et al. 2004). That is, aspects of our material environment can
impact our behavior. CV studies show that we use race and gender markers to evaluate the same work differently, with evaluators showing a marked tendency to give more positive reviews to work presumed to have been done by men or by white people.

This all can manifest in microbehaviors, such as with whom we make eye contact, to whom we stand more closely, and the motivations we ascribe to people’s behaviors. All of this, in turn, can lead to increasingly large scale outcomes: who gets a chance to speak, who gets the job, or who gets the benefit of the doubt in potentially violent situations. So, to sum up, implicit associations can impact our perceptions of and interactions with others as well as our self-conceptions and performance at tasks.

It is well recognized that implicit associations are an important cognitive mechanism for our functioning in the world; they enable us to quickly call forth relevant information without having to first consciously sort through every belief we hold (Staats, et al. 2016). Without this, we would be cognitively overloaded and unable to act. Many implicit associations are helpful: associating lightning with thunder helps us quickly get to safety during a storm (Staats, et al. 2016). Yet not all are helpful, and many are actively harmful. For simplicity, I am going to use the term *pernicious implicit associations* to label implicit associations based on identity markers where the associations draw on, maintain, or establish oppressive or subordinating norms or social structures. Most obviously this encompasses negatively valenced associations regardless of their content (associations with negative words) and negative stereotypes. It also includes supposedly positive stereotypic associations, as well as representations dehumanized associations or representations that result in increased avoidance behaviors. These associations uphold a subordinating and oppressive status quos
and help to objectify and obscure the personhood of members of the social group in question.

Which implicit associations we hold matters. But because they are non-conscious and unreflective, it is extraordinarily difficult to identify what our implicit associations are. Once we know about which implicit associations we hold, we have limited options for dealing with them. If our implicit associations diverge from our reflectively held beliefs, as many of ours do, noting the associations and our disagreement with it simply does not do much to dispel the association (Staats, et al. 2016). Options to combat pernicious implicit associations include: quickly exorcising the association when we encounter it, limiting the effects of the associations we hold, or fostering alternative associations. Quickly exorcising the association when we encounter it can work well to block the pernicious association, but doing this requires that we be aware of the way we are being primed and that we have the cognitive resources to exorcise the association; doing this when we are under high cognitive load, such as when we are stressed or multitasking, is difficult (Huebner 2009). If we cannot block the association, we can work to counter its effects, say by anonymizing CVs from job applicants or by having students fill out demographic information after they have taken a test rather than before it. However, the most promising, if labor intensive, strategy is to foster alternate sets of associations in place of the pernicious ones. Focusing on a counter-stereotypical member of the stigmatized group has been found to help shift the valence of associations about the group overall (Staats, et al. 2016). The best way to shift pernicious associations, though, is to interact with a variety of people from the targeted group so as to dismantle monolithic stereotypic associations about the group (Staats, et al. 2016).
There is a lot of ongoing debate about whether and how we might be morally responsible for the implicit associations we hold (Zheng 2016). What is clear now, though, is that fostering pernicious associations in ourselves and others is harmful; it negatively impacts the subject’s capacity to act as a moral and epistemic agent (Gendler 2011). Given that it is harmful, prima facie we ought to avoid doing it and, if we cannot avoid it, we ought to be particularly careful and responsible in when and how we engage in an action that will call forth these associations.

IV. Slurs as priming words

There is good reason to believe that slurs can act as powerful priming mechanisms for pernicious implicit associations. While this has not received extensive study, the work that has been done paints a compelling picture. Greenberg and Pyszczynski (1985) found that racial slurs activate negative schema of the targeted group more so than general negative criticism, especially when the target’s behavior conforms with stereotypes about the group.\(^{19}\) Kirkland, Greenberg, and Pyszczynski (1987) found that the negative schema activated by the racial slur spilled over onto evaluations of individuals connected to the target of the

---

\(^{19}\) Greenberg & Pyszczynski (1985) examined the effect of racial slurs on evaluations of skill. White participants observed as a black debater either won or lost a debate. Participants then overheard criticism with a racial slur, criticism with no racial slur, or no criticism. In cases where the black debater lost, participants who overheard the racial slur rated the debater’s skill more negatively. No effect was found when the comments did not contain the racial slur or when the black debater won. This study found that slurs activate negative schema of the targeted group more so than general negative criticism, especially when the target’s behavior conforms with stereotypes about the group.
slur. Carnaghi and Maass (2008) found that slurs versus category labels elicit more negatively valenced implicit associations, increased avoidance of those targeted by the term, and increased intergroup, rather than interpersonal, thinking. Fasoli et al (2015) found that sexual slurs foster significantly increased dehumanization and avoidance behaviors of the targeted group in contrast to category labels and offensive general insults. Taken together

---

20 Kirkland, Greenberg, and Pyszczynski (1987) examined how the negative effect of overheard slurs extends outward from the target. White participants read a trial transcript with a white defendant and either black or white defense attorney. When the attorney was black, participants overheard a comment referring to the attorney as either a “shyster” or a racial slur, or no comment at all. The racial slur led to derogatory evaluations of the attorney similar to the findings in Greenberg and Pyszczynski (1985). The defendant received especially negative evaluations and harsh verdicts when defended by the black attorney who had been targeted by the racial slur. The study found that the negative schema activated by the slur spilled over onto evaluations of individuals connected to the target of the slur.

21 Carnaghi and Maass (2008) explored the implicit associations that arose in response to exposure to slurs versus category labels. The findings show that slurs prime audience members with more negatively valenced implicit associations than do category labels. While there did not appear to be any difference in terms of the stereotypic content of an audience’s implicit associations about members of the group, exposure to the slur made the audience more likely to associate group members with negative terms. Exposure to the slur led to a greater facility for negative words, while exposure to the category label led to a greater facility for positive words. Those exposed to the category label were faster to approach members of the targeted group than those who were exposed to the slur. Finally, slurs prime hearers to understand the situation in terms of intergroup interactions, rather than in terms of interpersonal interactions. That is, slurs prime hearers to understand a situation in terms of in-groups and out-groups, ‘us vs. them,’ rather than in terms of individual people interacting with one another.

22 Fasoli et al (2015) examined the effects of slurs contrasted against insults and category labels. Participants were exposed to a sexual slur, corresponding category label, or offensive though non-slurring insult (i.e., ‘asshole’) and then were evaluated using the IAT. Exposure to the slur fostered dehumanization of the targeted group via a significant decrease in human-related associations about the targeted group, in contrast to associations after exposure to the category label or the insult. This result held for both supraliminally and subliminally priming. Subliminal exposure to the slur also led to increased avoidance of members of the targeted group (avoidance was not evaluated in the supraliminal priming study.)

23 See also: Henry, Butler and Brandt (2014); Gaucher, Hunt and Sinclair (2015); Carnaghi and Maass (2007); O’Dea, et al. (2015). Within philosophy literature, Camp’s *perspectivalism* is
and put simply, slurs powerfully prime pernicious implicit associations about the targeted group. Note that these results do not indicate that category labels don’t have any priming effect on audiences. Rather, they indicate that slurs are particularly powerful in calling forth pernicious associations about members of the group.

So far, this gives us good reason to believe that slurs can prime pernicious associations, but it doesn’t yet indicate whether this results from using the slur, mentioning it, or from some other mechanism. While the use/mention distinction is usually difficult to map onto psych studies (reading a vignette does not easily classify as exposure to either a use or a mention), several of these studies have addressed this issue directly. Greenberg and Pyszczynski (1985) specifically looked as uses of the slur and (unsurprisingly) found that this called forth pernicious associations. More strikingly, at least two of these studies give strong reason to believe that these pernicious implicit effects result from mere exposure to the slur, rather than only to the use of the slur. Carnaghi and Maass (2008) exposed the audience to the slur and category label, as a single printed word, and then tested audience implicit responses. Fasoli et al (2015) examined both simple exposure, consistent with Cargnaghi and Maass’s methodology, and subliminal exposure of 13 milliseconds to a slur, category label, or offensive general insult and found consistent pernicious effects from both forms of exposure. In both studies, pernicious implicit responses arose from the mere exposure to the term, rather than from a contextual use of the term.24 Indeed, this is hardly surprising given

quite close to this view, though Camp stops short of explicitly focusing on the role of implicit associations; Camp (2013).

24 Some people have raised the concern that in both forms of exposure, the audience is tacitly understanding the term as coming from a speaker and this is the real source of the associations primed by the word. It’s unclear from the evidence currently available if this is the case, though nothing seems to indicate that this is what is going on. Even if it is, this
what we know about these processes: there is little evidence to suggest that our low-level (implicit) processes are able to differentiate between the use and the mention of a slur. What this means is that both using and mentioning the slur can prime audience members in this harmful way. Merely mentioning a slur can act as a powerful priming mechanism of pernicious implicit associations.\textsuperscript{25}

These studies, for the most part, focus on the impact of slurs on outgroup members, i.e., those who are not targeted by the term. Though ingroup members were excluded from the studies, there is strong reason to believe that people targeted by the slur are also impacted, and more than that, are impacted in multiple ways. There is good reason to believe that ingroup members are impacted by 1) internalized pernicious associations, 2) outgrouping and exclusion, and 3) stereotype threat.\textsuperscript{26} First, it has been well documented that ingroup members internalize pernicious associations about the group of which they are a part (and sometimes do so even more strongly than others due to repeated exposure to the stereotype) (Greenwald and Krieger 2006; Rudman, Feinberg, & Fairchild 2002). There is little reason to think that this does not hold with slurs – though the details of the associations called forth does not change my argument: if audiences are automatically encoding even subliminal exposure of a slur as coming from and being endorsed by a speaker, then the same will be true in cases where a speaker is mentioning the slur and the same pernicious effects will result. If anything, this gives us further reason to tread carefully as it would show that presumed speaker-endorsement is so tightly connected to how we interpret the slur that it’s unavoidsable even when audiences are not aware of being exposed to the slur.

\textsuperscript{25} This result gives us reason to discard theories of slurs that locate the derogation fully within the semantics of the slur. On these theories, it is possible to mention the slur without causing offense or harm. Given what we have good reason to believe about the priming effects from mere exposure to the slur, then that stipulation doesn’t work – at least part of what’s going on with slurs is happening in the pragmatics.

\textsuperscript{26} Ingroup members are also impacted by the behaviors of others who are primed by the slur, as well as by the myriad ways the presence of slurs shapes the pragmatics of the interaction.
ought to be studied. So it is likely that people targeted by the slur experience at least the same pernicious effects as do those who are not targeted by the slur. However, it is quite likely that the harm done to those targeted by the slur goes much further, via outgrouping and stereotype threat.

Slurs perpetuate outgrouping and exclusion. Carnaghi and Maass (2008) found that slurs activate intergroup thinking, so that those exposed to the slur tend to understand the situation in terms of ‘us vs. them.’ The impact of group identity representations on those who are outgrouped by such representations has been well studied (Carnaghi and Maass 2007; Master, Cheryan, and Meltzoff 2016). Representations that signal this kind of ingrouping and outgrouping lead to increased feelings of exclusion amongst those who are outgrouped (Schmitt, et al. 2010). Schmitt et al. explored how signifiers of ingroup membership can lead to feelings of alienation and exclusion in those who are not members of the group. Schmitt et al. studied the impact of Christmas decorations on those whose identities are consistent with celebrating Christmas and those whose are not. Though participants in both groups predicted that Christmas decorations would not negatively impact their mood, the study found that such decorations significantly impacted participants’ feelings of inclusion and exclusion. Specifically, those who were outgrouped by the representation experienced decreased self-assuredness and self-esteem, both of which were mediated by increased feelings of exclusion. Given that slurs activate intergroup thinking and that other, more accepted, representations that lead to such thinking produce significant feelings of exclusion, it is likely that slurs lead to at least similar, if not more powerful, feelings of exclusion in the targeted group.
Finally, ingroup members are impacted by the slur via the stereotype threat it calls forth. Stereotype threat arises when negative stereotypes about a stigmatized group are made salient such that an individual member of that group is concerned that their behavior will be taken to confirm the stereotype (Steele, Spencer and Aronson 2002). Multiple studies have found that when stereotype threat exists, individuals perform worse at tasks than they would under no-threat conditions (Beilock, Jellison, Rydell, McConnell, and Carr 2006; Schmader and Johns 2003; Spencer, Steele, and Quinn 1999; Steele and Aronson 1995). The stereotype can be made salient either blatantly or subtly, with different results stemming from each process (Stone and McWhinnie 2008). Explicitly cuing the stereotype causes targets to adopt a ‘prevention focus orientation’ in which they focus on keeping their behavior from conforming to the stereotype, which often has the ironic result of worsened performance at the task. Implicit stereotypic cues, on the other hand, are found to increase cognitive load in the target which decreases working memory-based performance, again leading to worsened performance at tasks. When both blatant and subtle cues occur together, targets’ behaviors are impacted through both processes (Stone and McWhinnie 2008).

Using slurs is generally recognized as a blatant way of making stereotypes about the target group salient and so causing stereotype threat in individual members. Given that exposure to a slur recruits pernicious implicit associations, there is good reason to believe that the mere presence of a slur is itself a subtle cue and is sufficient to elicit implicit stereotype threat in members of the targeted group. This means that when a slur is used it is plausible cueing stereotype threat along multiple dimensions. Moreover, even when an audience is merely exposed to a mention of the slur, members of the targeted group are likely to be working under an increased cognitive load that impacts their working memory
processes, such as retaining information, following the train of thought in a lecture, or evaluating an argument or experiment. This matches well with the reported experiences from people in the targeted group who have been exposed to mentions of slurs.

If it is the case that mentioning a slur can act as a powerful priming mechanism of pernicious implicit associations for both in-group and out-group members, this not only raises important questions for how we ought to talk about slurs, but it also (helpfully!) gives us information for going forward as well. If mentions of slurs can serve as powerful mechanism for priming pernicious implicit associations, then regardless of the way it is discursively embedded the slur can call forth these pernicious associations. This gives us very good reasons to approach discussions of slurs carefully. Even mentioning slurs carries with it the very real possibility of negatively priming audience members, which can impact their self-conception, performance at tasks, feelings of inclusion, and perceptions of and interactions with others. Mentioning slurs call forth implicit associations that draw on, maintain, and establish oppressive, subordinating, and dehumanizing norms based on social group membership. And, moreover, that slurs do this particularly powerfully. This not only raises important ethical questions for how we ought to talk about slurs, but it gives us helpful information going forward.

V. How to talk about slurs

For those of us whose job it is to work on slurs, we have strong reasons to proceed with caution. I encourage us to recognize that each time we deploy a slur, this action may constitute a harm to the audience. This does not necessarily mean we can never discuss any
slurs, but it does mean that we need to be much more cautious in doing so. In order to help us think about this, I am going to identify some of the issues we ought to be taking into consideration before mentioning slurs. I’m not going to offer a hard and firm principle we can follow to determine whether to utter a slur -- in part because this is so complicated and context-dependent. Rather, I’m going to identify the issues we ought to be responsive to in navigating how we talk about slurs.

What is the purpose?

Before anything else, we should think carefully about what our purpose is in mentioning a slur. Given that we have good reason to believe the mere presence of a slur can be harmful, and harmful to already vulnerable groups of people, we should consider who is benefiting from mentioning the slur. Generally, we oughtn’t mention slurs unless doing so is going to ameliorate the harms connected to the slur. This can be broadly construed – it can include working to understand how slurs perpetuate oppression, how they impact people, deconstructing those slurs, or even things like flagging for a new English speaker that the term is a slur and that it’s harmful to use. But while this can be broad, attending to this will rule out instances in which a slur would be mentioned, say, to shed light on a semantic puzzle purely for the purpose of understanding the puzzle, or to titillate people, or to simply to increase web traffic to an article.

Can mentions be avoided?

If we are engaged in such an ameliorative project, we still shouldn’t jump right in to mentioning slurs. If there are other ways to make the point that don’t involve deploying the
slur, we should choose that path. Common strategies to do this include talking about slurs generally rather than putting forward a specific slur, using placeholders or symbols, turning to historical slurs, or drawing from fictional slurs in pop culture. I’ll discuss each of these in turn.

If it does not actually matter which slur is under discussion, there is no need to mention a specific slur. Say, if someone is making the point that using slurs is a macro-aggression rather than a micro-aggression, it is probably unnecessary to mention a specific slur as an example – just identifying the category ‘slur’ is sufficient. Similarly, ethics of journalism articles discussing how to report on uses of slurs probably do not need to deploy specific slurs in their examples. (Note that this chapter does not deploy any slurs except insofar as they appear in citations, and the thesis overall mentions slurs sparingly.) While the push for specific examples is usually laudable, talking about slurs may be the exception. Here, it is often better to keep the discussion more general unless examination of particular slurs is specifically necessary.

Another strategy is to use symbols or placeholders for the slur. Sometimes these representations are decipherable as a particular slur, though they need not be. Philosophers often use placeholders such as ‘p’ in place of a particular slur. This works especially well for examining the role of a slur in an utterance. Other times, though, people use the popular strategy of referring to a slur by its first initial (the ‘Z-word’) or writing it as the first initial followed by the dashes or other symbols (‘Z---’ or ‘Z***’). Whether this works is an empirical question and, to date, the psychological effects of this strategy have not been studied. However, there are two possibilities that seem equally plausible to me. It may be that this strategy simply causes the audience to think the slur to themselves, which leads to
the same harmful effects as mentioning it. If that is the case, this is especially insidious as it
gives the veneer of the speaker having ‘clean hands’ while offloading the responsibility for
priming the pernicious associations to the audience members themselves (see Louis C.K.’s
[2010] excellent discussion of this.) Alternatively, this strategy might lead to importantly
different effects than mentioning the slur: by refusing to include the full word, the speaker is
flagging that the slur is taboo and wrong. This could work as a signal to the audience which
would better enable them to quickly exorcise the associations called forth by the slur. My
guess is that even if the latter result holds some of the time, it does not hold all the time.
Still, in those cases where it is necessary to identify a particular slur this may be a good
compromise. However, this warrants further study as it is likely the strategy does not work as
well one might assume.

Sometimes people opt to use historical slurs, (such as the outdated slur ‘bosch’ for
Germans) that have fallen into disuse such that there are not associations to the term any
longer (Anderson and Lepore [2013] adopt this strategy.) This works particularly well in
cases where the speaker is working to convey differences between slurs and category labels.
Indeed, this is one of the strategies that has been largely adopted within the philosophy
literature on slurs. However, specifically because the slurs here have fallen into disuse, they
won’t have the same emotional resonance with audiences as would ‘live’ slurs. They may not
serve as well at helping an audience to truly understand the slur’s role within systems of
oppression and dehumanization. When this is the purpose in discussing slurs, this strategy
has significant drawbacks.

A strategy that can work well in a classroom is to turn to pop culture, and draw from
well-known fictional examples of derogation. Within the fictional world of Harry Potter, for
example, wizards who are born in non-magical families are called Muggle-born. But there is also a deeply derogatory term which is used, ‘Mudblood’ – and the books and movies really explore the intense marginalization and oppression of this group. The reason this strategy can be effective is that many people are familiar with this piece of pop culture and can grasp the significance of this as a derogatory term. From this we can get the clarity and emotional pull of the example without priming associations that are going to impact how we interact with real people out and about in our daily lives. This works particularly well in instances where students are discussing how various theories apply to slurs, and the necessary repeated mentions of slurs come off as cavalier. Of course, fictional examples do not always work and depending on the context can trivialize the issue.

While all of these alternative strategies have strengths and drawbacks, sometimes it may be necessary to explicitly mention a slur. Even when a speaker is warranted in mentioning a slur, they ought to do so sparingly; just because it is for the right cause does not give a speaker free reign to deploy the slur as often as they wish. Studies have shown that even when an audience is aware that the information they are being exposed to is false or irrelevant, the mere repetition of exposure helps to foster an implicit association and makes the audience more likely to rely on this association in moments of high cognitive load (Gilbert, Tafarodi and Malone 1993). This means that repetition ought to be limited as much as possible. Again, this substantially reduces many of the mentions of slurs that we see going on in classrooms, in talks, in papers, and in journalism.
Who has social power?

Next, we should take into consideration the relevant power differentials amongst those part of the discursive interaction. If the speaker inhabits a position of greater social power in relation to the audience, the speaker ought to be more wary in deploying the slur. Social power can come from many, often interconnected, sources, such as an official role or a privileged social status due to race, class, gender, or other identity categories (and often from the intersections of these). The dimensions of social power amongst participants in the interaction matters for at least two separable, though in practice often interconnected, reasons.

First, when biased speech goes unchallenged, audiences are more likely to encode the content of the speech as less problematic than they previously reported such content to be (Rasinski, Geers and Czopp 2013). A person who in other contexts identifies, say, racist speech as problematic, who is exposed to such speech and does not take a structural opportunity to challenge the speech (which does not take into account social barriers), is more likely to downplay how problematic that instance of racist speech was. Yet people are not equally positioned to push back against harmful utterances: individuals with marginalized identities are less likely to enact resistance to such speech, often due to well-founded fear of repercussions (Nielsen 2012). So, when a speaker with a great deal of social authority mentions a slur to an audience with reduced social authority, the speaker is less likely to be challenged and so the audience is more likely to internalize the speech as less problematic. Given that the presence of the slur is perceived as less of a problem, audiences are then less likely to engage in ameliorative practices such as quickly exorcizing the associations called forth by the slur, instead letting the pernicious associations called forth remain unchecked.
This means that in cases of power differentials between the speaker and the audience, the presence of the slur may be more harmful than it would be otherwise.

Second, when the speaker is high status person and from a group not targeted by the slur, their identity combined with mention of the slur may trigger increased stereotype threat for in-group members. Multiple studies have found that when a person from a high status social group officiates over a task, participants who are susceptible to stereotype threat are likely to perform more poorly on the task then they would if the officiator were a member of a marginalized group (Danso and Esses 2011; Marx and Goff 2005). So, white men presiding over either women or black participants are more likely to trigger high level stereotype threat in potentially stereotype-laden contexts than are women or black officiators. When the speaker is saying a slur, this creates a stereotype-laden context. When the speaker is a from a high status social group and is not a target of the slur, they may foster increased stereotype threat amongst those targeted by the slur, which, again, has increased cognitive effects (Davis III, Aronson and Salinas 2006).

In practical terms, there is a morally salient difference between a professor mentioning a slur in a classroom and a student doing so – though, sometimes the microdynamics of student social life lead to students having a great deal of social authority. Similarly, a person with a great deal of social power, such as a white man, even if he is not the official authority figure in the interaction, ought to be more reticent in saying slurs. In both cases increased social authority, especially in contrast to the identities of audience members, can lead to greater harmful effects from the pernicious associations recruited by the slur. This means that when a person is in a position of power they may have a greater obligation to find alternate ways of making their point.
Who is the speaker?

In conjunction with this, understanding the connection between slurs and implicit associations can help us to understand how identity markers can impact the speaker’s standing to deploy a slur. There is a widespread sense that members of the group targeted by the slur seem to have a different standing when it comes to deploying the slur, though philosophers have struggled to explain why, exactly, this is. I suggest that one reason for this difference is that ingroup members may be doing different things with the associations they call forth.

Identity markers, such as skin pigmentation or appearing to have breasts, are well known priming mechanisms of implicit associations. When the speaker is a member of the targeted group, their identity, if it is apparent to the audience, can impact what they are doing in saying a slur. While even less research has been done here than on other aspects of slurs, there is some reason to believe that ingroup speakers are positioned differently than outgroup speakers with respect to the effects of saying a slur.

First, members of the group targeted by the slur may be able to say the term without fostering increased stereotype threat among other members of the targeted group in the way that high status outgroup members do. This is supported by the findings of Stone and McWhinnie (2006): Members of the group targeted by the slur may still be recruiting pernicious associations when they say the slur, but they may be able to avoid placing an increased cognitive burden on other ingroup members. Again, this tracks the reported experiences of ingroup members who are exposed to slurs issued by other ingroup members. Even if this is the only difference identity makes, it is significant in explaining why ingroup members may be differently positioned to say the slur.
Second, the speaker’s identity, if it is perceivable to the audience, may serve to facilitate alternate sets of associations about the group. Though this is far from proven and warrants further study, it is suggested by the many studies that have shown that focusing on an admired ingroup member may alter the valence of the associations an audience member has about the group. This is further strengthened by the findings of Galinsky, et al. (2013) who found that speakers who self-label with slurs targeting a group of which they are a member are able to foster positive perceptions about the group. This puts the speaker in an importantly different position from one who is not a member of the targeted group. Rather than committing harm in service of dismantling mechanisms of oppression, an ingroup speaker may be able to foster positive associations amongst the audience. It may well be the case that an ingroup speaker is, in fact, doing different things with their words than is an outgroup speaker.

Of course, even if this hypothesis is right, it cannot be assured in any individual instance: while identity markers may be reliable priming mechanisms, using speaker identity to foster positive associations is not guaranteed to work. This is especially so amongst already high-prejudiced audiences, who are likely to take the speaker’s identity as reason to dismiss or devalue their authority and whose implicit associations are more entrenched and resistant to change (Carnaghi and Maass 2007). In cases where the speaker is not able to use their authority to trigger or foster positive implicit associations they may be committing the same harm as outgroup speakers (Herbert, Precarious projects: the performative structure of reclamation 2015). Nonetheless, even if only sometimes successful, if these two effects about

---

27 Galinsky, et al. (2013) found that speakers who self-label with slurs targeting a group of which they are a member are able to foster positive perceptions about the group.
ingroup speakers are correct then this provides new explanatory and moral insight into how social position impacts the contours of talking about slurs.

What is the discursive medium?

Next, we should take into consideration the medium of the discursive interaction; there are important differences between discussions occurring in real-time versus deferred interactions. This is an atypical way to carve up different discursive mediums: We are generally used to thinking about speaking in person vs. say, through an article. But verbal vs. written does not properly track the salient differences for my purposes here. Rather, what matters here is the time lag between when an audience receives the utterance and when they are expected to process and respond to it. Real-time interactions refer to back-and-forth, in the moment conversation, whereas deferred interactions are when the audience has time to process and reflect on the discursive contribution without this kind of time pressure.

What is important is how the medium impacts how we process and respond to information. People reading an article at their own pace can usually take the time to question and exorcize negative associations as they are encountering them. Audience members encountering the slur in real-time, though, are under an increased cognitive load: they must follow the conversation, process the argument being given, and attend to all the other subtle features of the situation without the luxury of slowing down and mitigating the effects of the slur. That is, there is a real, ethical difference between deploying slurs in real-time discursive contexts and in contexts where the audience can more effectively deploy strategies to mitigate the pernicious associations called forth by the term. Put simply, it may be more
permissible to mention slurs in, say, articles, than in talks, in a classroom, or on discussion forums.

The upshot of this is that we need to attend carefully to the dynamics of the interaction in which the slur is mentioned, rather than only to whether it is spoken or written. This means that the practice of showing text of the slur, rather than saying it, during a presentation may not work as well as those who use this strategy might hope. This also raises complications for all the kinds of discursive mediums where it is unclear whether they are occurring in real-time or with a time delay: things like texting, messaging, or commenting on blogs and forums. Sometimes these interactions are obviously occurring as real-time exchanges. But sometimes, someone else - another commenter perhaps - can encounter the exchange and read it as a static document. Further, these static documents can persist far into the future and may be read or listened to by a wide array of audiences. While it makes sense to err on the side of caution in issuing slurs, this also points to an obligation to be sensitive to the evolving mediums by which we communicate.

**Can harm be reduced?**

Being responsive to these issues sets a high bar for when it is acceptable to utter a slur. On occasion, there is sufficient reason to do so. When we've decided it’s necessary to utter a slur, we ought to think about how to diminish the distress inflicted by the slur. One way to do this is by preceding the discussion with a content warning. Content warnings are ways to alert the audience to the content of a show, text, or discussion before they are exposed to it. Content warnings are the broader category into which trigger warnings fit (Smith 2014). While trigger warnings are specifically warnings to survivors of trauma that the upcoming
content may call forth psychological and physiological responses from that trauma, content warnings are more general. Alerting an audience to an upcoming slur serves several productive purposes. Content warnings are a way to flag that slurs are not socially or morally acceptable. While in most instances saying a term serves as a way to license others’ use of that term, content warnings can be a way to expressly mark that this is not the case here. While hopefully everyone already knows this, it is important to make it explicit – especially, this is important so that the speaker is not simply complicit in licensing the slur. By expressly establishing for the audience that the upcoming term is problematic, this can also serve as a recognition that the slur may differently impact various participants – those who are targeted by the slur are forewarned about where the conversation is going and can make informed choices about what they enter into. A person who lives their daily life within a racially biased social world may, reasonably, decide that they are simply too exhausted on a particular day to expose themself to racial slurs.

Content warnings can serve a particularly important role in the case of talking about slurs. Flagging an upcoming slur can decrease cognitive load when the slur arrives. Audiences then don’t have to process that a slur is being said at the same time as it is being said, and instead can focus on other things. This better prepares the audience to quickly exorcise the implicit associations the slur calls forth. By forewarning audiences that a slur is going to be issued and that this slur is problematic, the audience is better positioned to mitigate the pernicious effects of the slur.
VI. Revisiting puzzles about slurs

If all of this is right it can help to shed light on the central puzzles about slurs I laid out earlier. To begin, one of the things slurs do is to serve as particularly powerful priming mechanisms of pernicious implicit associations. This position isn’t wholly unfamiliar within philosophy. Elisabeth Camp’s *perspectivalism* and Eric Swanson’s account of the connection between slurs and ideology are both compatible, and at the very least make gestures towards, the ways in which slurs can prime implicit biases. Camp argues that slurs call forth and license a particularly hateful perspective, and so even when we deny the application of a slur, simply using it still calls forth this perspective (Camp 2013). Swanson argues that slurs and ideologies mutually support each other: slurs elicit or cue particular ideologies, and at the same time using a slur makes the ideology to which it is connected more acceptable (Swanson 2016). I take these views to be generally compatible with one another, though they’re highlighting different important aspects of slurs. However, neither Swanson nor Camp delve into just what it means to elicit an ideology or to call forth a perspective. Both views leave open that at least part of what is happening is within implicit, low-level processes. My view offers a way to fill in the precise connection between slurs and implicit associations in order to fully make sense of their views. Furthermore, filling in these details better helps us understand just what it is that slurs do, and how they do this.

Beyond helping to make sense of existing views on slurs, drawing out the way slurs prime pernicious implicit associations sheds light on the pragmatics of slurs. To date, much of literature on the pragmatics of slurs has focused on illocutionary force. While this is one important aspect of slurs (which I discussed in Chapter 3, “Derogative: Insults, Slurs, and
Speech Acts”), the perlocutionary effects of slurs has gone largely unexamined. Priming pernicious implicit associations is a significant effect of slurs. That they do this particularly powerfully is part of what differentiates slurs from category labels.

This also helps to make sense of why denying slurs doesn’t seem to work easily – these pernicious priming effects occur even when a speaker is denying the application of a slur. This is at least part of why denying the application of slurs seems to have some sort of spillover – because the priming effects hold regardless of whether the application is being affirmed or denied. Attending to the psychological response queued by the slur may also help to make sense of the real difference that the speaker’s identity makes. Similarly, it seems likely that a substantial part of successful reclamation is centered on changing the associations primed by the slur – either changing the power of the slur to elicit those effects or changing the sets of associations called forth in the first place. And, finally, this account of slurs sheds light on why use/mention seems to break down in the case of slurs. This gives us reason to engage with slurs reflectively, both in terms of reflecting on what slurs do and how they are used and in terms of navigating the contours of talking about slurs.
BIBLIOGRAPHY


http://www.huffingtonpost.com/entry/unc-sexual-assault-accusation_us_57d83eece4b09d7a6880018c.


National Congress of American Indians. 2013. "Ending the legacy of racism in sports and the era of harmful 'Indian' sports mascots."


Rose, Diana, Graham Thornicroft, Vanessa Pinfold, and Aliya Kassam. 2007. "250 labels used to stigmatise people with mental illness." *BMC Health Services Research* 7 (97).


