

SUBORDINATING SPEECH AND THE CONSTRUCTION OF SOCIAL HIERARCHIES

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

Michael Randall Barnes, M.A.

Washington, D.C.
August 21, 2019

Copyright 2019 by Michael Randall Barnes
All Rights Reserved

SUBORDINATING SPEECH AND THE CONSTRUCTION OF SOCIAL HIERARCHIES

Michael Randall Barnes, M.A.

Thesis Advisor: Rebecca Kukla, Ph.D.

ABSTRACT

This dissertation aims to deepen and broaden the philosophical literature on subordinating speech. It aims to demonstrate that *how* language subordinates is more complex than has been described by most philosophers writing on this topic. By departing from the rigid confines of the standard Austinean picture, and by drawing on recent work in oppression theory and political philosophy, I develop a more expansive conception of subordinating speech that more accurately captures lived realities of this problem.

I argue that the philosophical analysis of subordinating speech as it currently exists is limited in several ways, and that these limitations amount to an impoverished understanding of what are fundamental aspects of these speech acts. The harms that subordinating speech inflicts on its targets (Chapter One), the type of authority that is exercised by subordinating speakers (Chapters Two and Three), and the expansive variety of subordinating speech acts themselves (Chapter Three) are all under-developed areas in need of further refinement—and, in some cases, large paradigm shifts. Other topics like the explosion of abusive speech online (Chapter Four) or the distinctively collective speech acts of protest groups (Chapter Five) have yet to be adequately addressed by philosophers, and I argue that these need to be considered alongside the ‘paradigm’ cases of subordinating speech that inform most models.

Throughout, the importance of speaker authority is addressed and evaluated. Instead of seeing this authority as reducible to either a formal position or a merely local, linguistic phenomenon, I argue for a conception of speaker authority that is a richly contextual social fact,

distributed unevenly among members of different social groups. In particular, I develop an account of collective authority that explains how many speakers can join together to subordinate in a way that no individual speaker is capable of doing. This account, I claim, is better able to explain the social reality of subordinating speech than individualist models. With a more fine-grained account of subordinating speaker authority, I claim we are led to a more accurate picture of the different subordinating speech acts available to different speakers, along with how these may harm their targets.

ACKNOWLEDGEMENTS

This project is dedicated to my mom, who has always been amazing even as she faced significant health challenges, and my dad, who has been by her side throughout it all. Thank you both. My brothers, Matt and Tom, are also pretty good.

The actual content of this dissertation was only possible with the support and mentorship of my committee: Rebecca Kukla, Alisa Carse, Madison Powers, and Lynne Tirrell. I have learned so much from your work and from our conversations, which I hope continue into the future. I feel extraordinarily lucky to have landed at Georgetown University where the faculty's commitment to pluralism and engaged philosophy is something I wish every philosophy student could experience.

The friends I made at during my studies at Georgetown improved every aspect of my time here, including this dissertation. To Nabina Liebow, Hailey Huget, Matt Shields, Joey Jebari, Keyvan Shafiei, Megan Dean, Katherine Ward, Gerald Taylor, Omar Talhouk, Cassie Herbert, and others, I thank you so much for the time we got to spend together. I could not imagine a greater group of friends with whom to do philosophy—and do things besides philosophy too!

I also want to thank the staff at Georgetown, namely Carina Olsson, Nicole Richardson, and Rachel Rubin, who solved the many unexpected problems (both technological and bureaucratic) grad school throws at you. Like I say in my emails: As always, thank you so much.

This project also benefited from the many commentators and audience members at all the conferences where I presented early versions of this work. The same should be said of my many students at Georgetown, who also served as testing ground for many of the ideas found here.

I'd also like to thank the impromptu philosophical community I found when I returned to Toronto to finish this project. Lisa McKeown, Emma McClure, Mercy Corredor, Federica Berdini, Dan Barron, Mark Fortney, and others helped me find community after I'd left the one I had at Georgetown.

Lastly, my partner Lara Small deserves the greatest appreciation possible. You are a singularly impressive person who improves every aspect of my life. You inspire me to do good, meaningful work, and I hope you like this.

Many thanks,
Michael Randall Barnes

TABLE OF CONTENTS

INTRODUCTION.....	1
1. OVERVIEW OF THE DISSERTATION	1
2. SPEECH AND HARM: DISMISSING SOME NAÏVE VIEWS.....	7
3. THE CORE CLAIMS OF SUBORDINATING SPEECH	12
CHAPTER ONE: THE HARMS OF SUBORDINATING SPEECH.....	14
1. INTRODUCTION.....	14
2. THREE AREAS OF HARM	24
3. AMBIGUITIES IN THE ILLOCUTIONARY	57
4. CONCLUSION	80
CHAPTER TWO: SPEAKING WITH (SUBORDINATING) AUTHORITY.....	82
1. INTRODUCTION.....	82
2. SUBORDINATING SPEECH AND THE AUTHORITY PROBLEM	85
3. MAITRA’S SOLUTION TO THE AUTHORITY PROBLEM.....	88
4. PROBLEMS WITH MAITRA’S SOLUTION	91
5. INFORMAL SITUATIONAL AUTHORITY	97
6. CONCLUSION	104
CHAPTER THREE: PLURALISM IN SUBORDINATING SPEECH ACTS	106
1. INTRODUCTION.....	106
2. VARIETY AMONG SUBORDINATING SPEECH	115
3. A (ROUGH) TAXONOMY OF SUBORDINATING SPEECH ACTS	119
4. A HIERARCHY OF SUBORDINATING SPEECH	145
5. CONCLUSION	163
CHAPTER FOUR: WHOM DO YOU SPEAK FOR? AND HOW? ONLINE ABUSE AND COLLECTIVE SPEECH ACTS.....	164
1. INTRODUCTION.....	164
2. BACKGROUND: THE INTERNET AND ITS PROBLEMS	166
3. ONLINE ABUSE.....	173
4. ONLINE ABUSE AND THE CONSTRUCTION OF COLLECTIVE AUTHORITY	181
5. CONCLUSION	196
CHAPTER FIVE: THE PRAGMATICS OF PROTEST AND POSITIVE PROPAGANDA.....	199
1. INTRODUCTION.....	199
2. PROPAGANDA: POSITIVE AND NEGATIVE	201
3. THE MORAL-EPISTEMOLOGICAL FUNCTION OF SOCIAL MOVEMENTS	213
4. PROTEST AS PROVOCATION	220
5. CONCLUSION	224
REFERENCES	226

INTRODUCTION

1. OVERVIEW OF THE DISSERTATION

In this dissertation, I defend the claim that speech and language play significant roles in constructing and maintaining oppressive relations between social groups. In other words, I argue that things like hate speech, slurs, propaganda, abusive and discriminatory language, and more are important tools in the creation and sustenance of unjust conditions of domination and subordination in the broader society.

This, however, is a familiar idea for many—especially for members of the marginalized groups targeted by such speech. The contribution I aim to make in this project is to demonstrate that *how* language subordinates is more complex than has so far been described by most philosophers of language. That is, the philosophical analysis of subordinating speech as it currently exists is limited in several ways, and these limitations amount to an impoverished understanding of what I claim are fundamental aspects of these speech acts.¹ The harms subordinating speech can inflict on its targets (Chapter One), the types of power or authority that are exercised by subordinating speakers (Chapters Two and Three), and the variety of subordinating speech acts themselves (Chapter Three) are all under-developed areas within this domain of philosophy in need of further refinement, and in some cases large paradigm shifts. Other topics like the recent explosion of online abusive speech (Chapter Four) or the relationship

¹ Note that throughout I use the terms ‘oppressive speech’ and ‘subordinating speech’ interchangeably, simply for stylistic reasons. Some authors, like McGowan, tend to prefer “oppressive speech,” while others, like Langton, and Maitra, use “subordinating speech.” I am agnostic as to which term is preferable, though I default to “subordinating speech” more often.

between subordinating speech and other forms of socially-located speech acts—and in particular, protest (Chapter Five)—have yet to be adequately addressed by philosophers.

In this brief introduction, I first give a quick overview of each chapter’s main claims. I then go on to clarify my starting points by considering and dismissing some intuitive, but naïve views about subordinating speech. And finally, I situate myself within the existing literature, and describe what the main contributions of this project are.

Chapter One, “The Harms of Subordinating Speech,” states and explains my main thesis—namely, that speech harms. It does so by detailing the varied harms of subordinating speech identified in the existing literature, and then arguing that a focus on illocutionary acts—the focus favored by many leading theorists writing in this area, who have built upon the framework originally sketched by J. L. Austin (1962)—fails to capture central features of these harms. In brief, because the model of illocutionary acts focuses on discrete events—that is, singular transactions—it fails to offer a plausible and comprehensive account of how relations of oppression should be characterized in these speech acts. This is the case because conditions of oppression and subordination are essentially structural and for this reason cannot be captured in transactional terms. Moreover, the illocutionary act model for subordinating speech relies on the problematic notion of “paradigmatic” acts of subordination, which often include the intent to subordinate. This is something theories of oppression—and theories of oppressive speech too—can and should omit. The task of Chapter One, therefore, is to thoroughly examine the harms of oppressive speech as they have been discussed by various theorists, and to argue that illocution-focused accounts aren’t up to the task of adequately describing these harms.

In Chapter Two, “Speaking with (Subordinating) Authority,” I turn my focus from the targets of subordinating speech to its speakers and address how it is they—or, to be precise, some

of these speakers—have the power to effectively subordinate with their words. By taking on the so-called Authority Problem, this chapter aims to show how even ‘ordinary’ speakers are capable of marshalling strong social forces with their utterances. The Authority Problem—a recurring objection leveled at theories of subordinating speech—questions how ‘ordinary hate speakers,’ that is, your everyday person on the street, could possibly subordinate with their speech since the speech acts characteristic of subordination—ranking, authorizing, etc.—are standardly conceived as authoritative speech acts. I argue that while authority is in fact necessary for certain types of subordinating speech acts, the type of speaker authority needed here is a much broader category than is typically acknowledged. It can be informal, situation-specific, even ad hoc, while nonetheless destructive. This chapter looks to demonstrate *one way* in which speakers can come to achieve the sort of authority I conceive as relevant to subordinating speech, and in doing so sets the stage for the following chapters that go on to describe further types of subordinating authority.

Chapter Three, “Pluralism in Subordinating Speech Acts,” expands on the results of the previous chapter and argues for a rich variety among of the types of speaker authority that enable subordinating speech. It also explores non-authoritative entitlements that play a role in oppressive speech. In doing so, it expands on the taxonomy of types of subordinating speech acts as they are typically discussed, broadening our lens and making visible a variety of forms of oppressive speech. In doing so, I argue that current models conceive of the entitlements for subordinating speech far too narrowly, and as such fail to properly account for the speaker’s role here, by either overinflating or deflating it to fit a one-size-fits-all paradigm. I argue that by expanding our conceptions of the entitlements that enable subordinating speech and by

acknowledging the variety of types of subordinating speech acts, we're led to a more accurate account overall.

To see why, consider how the literature on slurs offers a good example of how productive it is to take note of the differences between subordinating speech acts. That is, it is a commonly recognized feature of slurs that they can differ along axes of offensiveness, strength, and so on.² As Mihaela Popa-Wyatt and Jeremy Wyatt note, philosophers often find slurs interesting because they demonstrate an “unusually rich set of effects” (2018, 2880). One pattern that is central to many accounts of slurs is the variable offense slurring speech acts generate.³ This refers to the fact that offence “varies across different slur words, across different uses of the same slur word, and across the reactions of different audience members” (ibid.). It is a common goal for many philosophers writing on slurs to develop a theory of slurs that accounts for this wide variety of slurring speech acts.⁴ Indeed, Popa-Wyatt and Wyatt (2018) argue that a necessary feature of a satisfactory account of slurs is that it accounts for these different offence patterns.⁵

Here, I'm less interested in how well different theories of slurs do in meeting this goal, and more concerned with suggesting that this starting point is one that the broader subordinating speech literature ought to take on board. Slurs, after all, are a common component of a lot of subordinating speech, and slurring speech acts are best understood as a subset within the

² The literature here is vast, and I largely leave it aside to focus on the subordinating speech act literature. But see Popo-Wyatt & Wyatt (2018); and Bolinger (2017) for recent overviews of some of the major positions.

³ Note that they distinguish between offence and derogation, where “derogation is part of the speaker’s communicative intentions, while offence is an achieved effect on the audience members, determined in part by their beliefs and values” (2018, 2881). I don’t agree with this way of distinguishing between these two concepts, but I put this aside for the purposes of a more concise exposition.

⁴ See, for example, Hom (2008; 2012); Anderson and Lepore (2013); Jeshion (2013); Whiting (2013); Ashwell (2016); Kukla (2018); Anderson (forthcoming); and Swanson (forthcoming).

⁵ This is because they take it that the goal of any account is to properly capture how these slurs function in the various ways they do. They go on to propose their own account: the role-power account, which makes use of McGowan’s conception of exercitives.

category of subordinating speech, which includes slurring acts and much more.⁶ That is, we should recognize that different *effects*—like distinct offence-generating patterns—are partly the result of different *speech acts*. Acknowledging this fact of variety—as obvious as it is—serves as an important first step and functions as a shared starting point for theorists writing about slurs that seems oddly ignored by dominant speech act accounts of subordinating speech. I believe philosophers interested in subordinating speech in general ought to follow this lead and recognize variety more explicitly, and I argue for this in Chapter Three.

Chapters Four and Five represent a slight change in focus. Rather than examine the ways in which the standard Austinean picture needs to be adapted—sometimes significantly—to better account for the realities of subordinating speech, these chapters build on the results of the earlier chapters and examines two areas that social philosophers of language have yet to examine in much detail. These are, first, the distinct forms of abusive speech that are found online (Chapter Four), and second, the interaction between progressive protest and reactionary counter-protest speech (Chapter Five). My goal in these chapters is to show both that these phenomena raise important questions worthy of philosophical investigation, and also that examining them leads us to a better understanding of subordinating speech in general. Chapters Four and Five are dedicated to what I call ‘collective speech acts,’ which I first introduce in Chapter Three. Both chapters look at how seemingly disparate utterances can combine to afford the speech a type of group-based authority, and in doing so offer powerful avenues for collective action.

Chapter Four, “Who Do You Speak For? And How: Online Abuse and Collective Speech Acts,” is about online abuse, and it has two main tasks. First, I argue this is a serious, but neglected, area of subordinating speech, and that social philosophers of language ought to pay

⁶ For an analysis of slurs specifically in terms of slurring acts, rather than slur terms, see Kukla (2018).

more attention to specifically online discourse. Second, I argue that accounting for the realities of subordinating speech online shows that speaker authority is dynamic and emergent, and often depends on the community in more ways than current theories of licensing and accommodation are equipped to explain. I consider how features like anonymity and the shared language of online abusers help to construct a type of group-based authority that lends potentially harmful power to their words. That is, rather than seeing these as a collection of disparate utterances, we do better to see online abuse as a collective activity, constituted by a type of collective speech act. I argue this better reflects the experience of the targets of online abuse and uncover the pragmatic reasons why.

I also suggest that a similar framework is appropriate for some forms of offline hate speech like racist vandalism and large hate marches, and so argue that online subordinating speech is not wholly different than its offline counterpart, but rather that the affordances of online communication platforms simply makes explicit what is often implicit in real life. It is in this way that online abuse represents, in some ways, a distinct phenomenon but also one that, at the same time, sheds light on IRL (in real life) subordinating speech. Overall my hope is to show that attention to online subordinating speech is useful for illuminating both the harmfulness of that phenomena itself, but also for clarifying features of IRL speech that regularly go unnoticed or under-emphasized in existing accounts.

Chapter Five, “The Pragmatics of Protest and Positive Propaganda,” builds on the previous two chapters’ conception of collective speech acts to consider the speech act of protest on the one hand, and the counter-protest speech it often generates on the other. I claim that, much like subordinating speech, protest is a clear example of socially-located speech, in part because an essential aspect of protest is its function of foregrounding the moral authority of the speaker.

This, in turn, can sometimes lead to counter-protest propaganda that aims to present the protesters as unreasonable, and therefore not worthy of serious consideration. I argue that this pragmatic-epistemic difference marks an important contrast between protest and propaganda. I argue that seeing each of these acts as collective speech acts, and furthermore as interacting with each other, helps to reveal their differences more effectively than does analyzing each independently of the other.

Overall, these five chapters aim to describe the reality of subordinating speech in all its complexity. It is by avoiding oversimplifications and fully recognizing the way things like hate speech, propaganda, abusive speech, and the like function in an interrelated way with so many non-speech oppressive forces that we begin to see the problem for what it is. While my main focus is on speech and language, throughout the dissertation I hope to make clear how these are inseparable from both the broader social forces that shape the meaning of our words, as well as the more visceral and violent events that grab people's attention.

2. SPEECH AND HARM: DISMISSING SOME NAÏVE VIEWS

Before I make a proper start in Chapter One, it's useful to acknowledge that some common ideas about speech and language, as well as the nature of oppression and harm, seem to present roadblocks to the general thesis mentioned above. It is best to take some of these on right away, if not to defeat or dismiss them once and for all, but to clarify my commitments from the beginning.

First, there is the issue of definition. It might be claimed that terms like 'hate speech' or 'propaganda' are simply too ambiguous, too slippery, or too ill-defined to afford any clarity on such contentious issues as harm and oppression. There is some truth to this, in that there are

numerous definitions on offer, and I share the skeptic's frustration that these are not always as precise as would be needed for some important tasks, like drafting legislation. However, I do not believe this problem is as unsurmountable as is sometimes claimed.

Since I am not here arguing for any kind of legal prohibition on hate speech—which would require careful precision in its definition—all that's needed for my purposes is simply a working characterization of the type of speech that can serve as a starting point in examining the harmful effects of subordinating speech. Hate speech, even roughly defined, serves that task here, offering a window to begin analyzing these harms, as well as those that less hateful though still subordinating utterances can inflict. So, while I do not deny the frustrating lack of a settled definition of 'hate speech,' I approach this topic by looking at the overlapping features that various definitions share. And though it is commonly frowned upon in philosophical circles, we can indeed start with the dictionary, which defines 'hate speech' as “[a]busive or threatening speech or writing that expresses prejudice against a particular group, especially on the basis of race, religion, or sexual orientation” (Oxford Dictionaries). Another definition appears in *The Price We Pay*, where Laura Lederer and Richard Delgado define hate speech as “speech or conduct aimed at a group of historically disenfranchised people; speech that reviles, ridicules or puts in an intensely negative light a person or group on account of who they are” (1995, 4–5). Finally, in the introduction to *Words That Wound*, Mari Matsuda, Charles Lawrence, Richard Delgado, and Kimberlé Crenshaw describe their focus on “assaultive speech,” that is, “words that are used as weapons to ambush, terrorize, wound, humiliate, and degrade” (1993, 1).

Taking these together, some immediately apparent common themes include the notion that this form of speech 'abuses,' 'reviles,' and 'degrades,' and in general expresses extreme prejudice for its target. These intuitive understandings, along with the importance of this speech

targeting historically marginalized groups, provide the necessary starting points for my account of the harms of subordinating speech. In what follows, I will clarify my focus as I move along, making use of the philosophical distinctions as I introduce them. Before that, however, some more common misconceptions must be (temporarily) dispelled.

Next, there is the widely held belief that speech, and communication in general, is simply the transmission of ideas—call this the communicative model of speech. Seen in this way, it's a mistake at best—and thought policing at worst—to suggest that speech can in and of itself harm someone. According to this view, even if I wanted to hurt someone with my speech, I simply couldn't do it—the old playground refrain of “sticks and stones . . .” is built upon this folk-philosophical foundation.

And yet, I believe we're all familiar with cases where speech does harm. Being subjected to a hostile and threatening rant can lead to increased anxiety and intense feelings of vulnerability and fear. Spreading (false) rumors about someone can lead them to lose their job or housing. Repeatedly being told that one's disability—or race, or gender, etc.—makes their life less valuable can contribute to a sense of inferiority, preventing one from pursuing important goals. Feeling threatened, having increased anxiety, losing a job or an apartment, and feeling constrained by other's expectations are all plausible harms, in that they inhibit a person's full flourishing. And they are harms whose source here is found in the speech of others. Some harms might be more immediate—feeling fearful and threatened in that moment—and others are more cumulative, the result of too many instances to keep track of, as is the case with constant small reminders of one's ‘inferiority’ or ‘deviance.’ And so, while “sticks and stones . . .” may be a memorable refrain, its errors are obvious. Indeed, it's worth noting that parents, guardians, and teachers often reach for the “sticks and stones” line precisely when words have clearly hurt.

A related but distinct common view is the while speech may, in some rare cases, amount to a harm, it never amounts to *oppression*. That is so, the thought goes, because oppression is a serious and grave wrong, usually taking the form of physical violence, and often backed by (unjust) law. The enslavement of millions of human beings was/is oppressive. Radically unequal marriage contracts, such as those that permit spousal rape, were/are oppressive. Lynchings were/are oppressive. But words? Mere speech? Words and speech are never oppressive on this naïve conception. They lack some of the (unstated) essential features that make something oppressive—whether that is the element of physical force and coercion, or perhaps some sanction provided by the state. While it may be admitted that speech can harm, it's denied that speech can oppress.

This is a more significant challenge, but in no way an insurmountable one. I will say more on this topic in Chapter One, here it is enough to say that the objection above relies on a far too narrow understanding of oppression. While gulags, guillotines, and gas-chambers might be the images that come to mind when oppression is discussed for some, this in no way detracts from the fact that segregation, discrimination, marginalization, and more are best understood as oppressive conditions for many social groups. Physical violence, though an important aspect of oppression, is not its sole expression. Attending to the roles language plays in supporting broader systems of subordination and oppression—which, it's worth noting, often still include instances of violence—is the goal of this dissertation. As I argue throughout, once we see subordinating speech as one aspect of interrelated systems of oppression, we should have no trouble understanding certain forms of speech as themselves subordinating acts. Furthermore, I would caution anyone from concluding that contemporary society lacks oppressive violence because of the (supposed) absence of the most extreme versions of this violence. As Patricia Hill Collins

notes, “the very pervasiveness of violence can lead to its invisibility” (1998, 66). Mass incarceration is a contemporary phenomenon, and a violent one too; racist policing is violent; and the extreme number of missing and murdered Indigenous women and girls vividly shows the indifference to violence that subordinated people routinely face. These conditions, and others like them, are supported and exacerbated by hateful and dismissive language. Thus, there is nonetheless a connection between speech and violent oppression, or so I will argue in more detail below.

One final initial roadblock I would like to discuss here is the belief that while speech can harm, and can indeed oppress, this only occurs when the individual speaker *intends* to do so. This may occur a lot or only a little, but given a putative example of speech that harms or oppresses, our main task should be to investigate whether that was the speaker’s intended effect. And where it wasn’t—and we’ll know this because they’ll honestly tell us they didn’t intend to hurt anyone—the insistence on moral blame is misguided and inappropriate. People, the thought goes, shouldn’t be held liable for the unintended effects of their speech; and they certainly shouldn’t be called ‘oppressors’ because someone else misunderstood what they meant.

The belief that speech is fundamentally about intentions, or the closely related belief that a speaker is only morally responsible for how they intended their words to be interpreted, is very common. And yet here again, there are relatable cases that betray this seemingly absolutist position. Even if the catcalling man truthfully reports that he was only intending to make a joke or pay a compliment, that does nothing at all to change the fact that many women interpret this behavior as aggressive and threatening. The comedian likely only intends to make people laugh when they joke about topics like rape and domestic abuse, and yet they are still blameworthy for the recklessness with which they can make some listeners relive traumatic memories. Failure to

see so much is more a failure to understand the criticisms being raised in these instances, than any sort of coherent view of blame. Negligence, even in the realm of speech, uncontroversially shows how the scope of morally blameworthy acts is wider than the scope of one's intentional acts. So we won't learn much if we only examine speaker's intentions in morally troubling cases.

3. THE CORE CLAIMS OF SUBORDINATING SPEECH

To quickly review the above, contemporary philosophers of language working on subordinating speech generally agree that (a) speech is about more than the communication of ideas, (b) speech can, in fact, harm, and (c) it can even oppress. Moreover, this can occur (d) even when it wasn't the speaker's intention to do so, and (e) in such cases they can rightly be held accountable for their (unintended) acts.

There are, of course, many disagreements within this sphere of agreement. Nonetheless, there is a consensus within this sub-discipline about what I will call the core claims of subordinating speech. These are: that subordinating speech acts are (1) possible, (2) prevalent, and (3) quite possibly fall outside the bounds of what is appropriately tolerated by the law and morality. My intention in this dissertation is to strengthen, complicate, and ultimately defend these core claims. I do so by drawing on recent work in social and political philosophy, and also by moving beyond the generally strict Austinean character of the existing accounts.⁷

⁷ While the views I go on to criticize are no doubt inspired by Austin, it is unclear how closely they follow the program as initially laid in by Austin himself, primarily in his *How to Do Things With Words* (1962). Austin no doubt took the illocutionary act—the act carried out *in saying* some utterance—to be his primary conceptual focus, but it is less certain that he found the perlocutionary effect of utterances—what occurs as a result of a speech act—to be as uninteresting or unanalyzable as those who followed him do. It is the overly *strict* or *narrow* focus on illocutionary acts that is the source of many of my critiques, and there is disagreement over whether this is attributable to Austin or not. For one recent account that pushes back against the narrow view as being Austin's that I largely agree with, see Nancy Bauer's *How to Do Things With Pornography* (2015).

Rather than abstraction, this project aims to be informed more by the lived reality of experiences of injustice and oppression. As I discuss in the coming chapters, many existing frameworks focus too intensely on individual speakers and their intentions, and as much of the recent—and not-so-recent—philosophical work on oppression makes clear, we should not limit oppressive acts to those of a specific agent who has a particular intention to act in an unjust way. Rather, we must attend to the structural relations that actions are bound up in and contribute to in order to properly describe subordinating speech. I argue that many existing accounts of subordinating speech fail to properly describe the relationship between individual subordinating speech acts and the broader oppressive norms that give these speech acts their force. The importance of speaker authority is central throughout and I develop a novel understanding of subordinating speech that re-conceptualizes the role of speaker authority in these acts. Instead of seeing this as reducible to either a formal position or a merely linguistic phenomenon, I conceive of authority as a richly contextual social fact that is distributed unevenly among social groups.

Overall, this dissertation develops a structural, pluralist account of subordinating speech. It is pluralist in the sense of recognizing the multiple avenues available for speech to act oppressively, and structural in that the primary focus is not on individual acts or intentions, but on the collective consequences of diffuse actions, some linguistic and some not. As I argue, these two features have not been adequately recognized by leading accounts in the existing literature.

CHAPTER ONE: THE HARMS OF SUBORDINATING SPEECH

1. INTRODUCTION

Philosophers and legal theorists who write on the potential harms of certain forms of speech tend to focus on explicitly hostile hate speech (e.g., “I’m gonna kill you f——t!”), assertions of genetic or cultural inferiority (e.g., “All Indians are savages.”), and the overt or covert call for genocide (e.g., “Muslims are an existential threat to civilization!”).⁸ These, along with the use of slurs, deeply entrenched epithets, and other cultural symbols of oppression like cross burnings and swastikas, are the paradigm cases of subordinating speech. Vivid examples like these—often real-life cases—are at the forefront of legal and philosophical analyses of the costs of, and plausible limits to, free expression.⁹

Words can hurt, as can other expressive symbols and gestures. Most theorists writing on the topic of speech and harm agree to this much. But there is less agreement about how to conceptualize the mechanisms of harmful speech and its damage. For legal theorists, the focus is mainly on speech as it functions as a form of assault—words that wound is the idea—and this naturally draws attention to the injuries suffered by the direct targets of such speech. For their part, philosophers working in the speech act tradition examine the *action* constituted by such

⁸ A note about slurs: throughout this dissertation I aim to avoid using or mentioning them as much as possible. While this may seem odd for a dissertation on subordinating speech, and other writers hold views about the necessity of confronting this language in all its ugliness, I personally do not wish to take part in their use. This stems partly from a disagreement over the possibility of any neutral use of such terms, along with the belief that a lot of productive work on this topic can be achieved while minimizing the reader’s exposure to slurs. For more on this, see Herbert (2018a).

⁹ For helpful overviews of some of the main issues and approaches to the topic of the harms of hate speech, see Matsuda, Lawrence III, Delgado, and Crenshaw (1993); Lederer and Delgado (1995); and Maitra and McGowan (2012).

utterances, conceived as a distinct matter from its contingent results, including the experience of injury that can vary from case to case.¹⁰

And while these are no doubt important aspects of the harm of such speech, they do not exhaust its harms.¹¹ This chapter aims to explore the harms of subordinating speech in a more expansive way, widening the lens of the philosophical analysis of oppressive speech in effort to better capture the wide range of harms attributable to such speech.

To that end, in what follows, I distinguish three areas of harm: (1) the negative psychological effects subordinating speech can have on its direct targets, (2) the harm constituted by the performance of the speech act itself apart from these contingent psychological effects, and (3) the broader societal damage done via subordinating speech, and in particular the maintenance of hierarchical relations of domination between social groups characteristic of oppression.¹² The main aim of this chapter is to catalogue and explain the harms of subordinating speech, which I do throughout section 2. Having done that, I go on to argue in section 3 that prominent accounts of oppressive speech—namely, illocution-focused models—fail to capture these harms adequately, primarily because they are centrally concerned with singular speech acts, rather than the broader *practices* of subordination that give these utterances their force.

¹⁰ Rae Langton's (1993) "Speech Acts and Unspeakable Acts," which put J. L. Austin's (1962) speech act theory to work in the service of identifying the subordinating and silencing nature of pornographer's 'speech' marks the beginning of this fruitful movement.

¹¹ Notable exceptions to the limited conception of harm I criticize below include Tirrell (2017); Gelber and McNamara (2016); and Ayala (2014). Both of which take a broader look at the harms of subordinating speech.

¹² In this chapter I note the different types of harms attributable to subordinating speech. In Chapter Three, I take a similar approach and argue that the different kinds of subordinating speech acts have also not been fully acknowledged by many philosophers writing on this topic.

1.1. Causing vs. Constituting

Before exploring the harms of subordinating speech directly, one key idea needs to be put in place. This is the distinction between speech (merely) *causing* subordination and some speech itself *constituting* subordination. This distinction is of central importance to the views I will go on to discuss. Understanding the distinction and the role it plays in those views is a necessary first step in evaluating how well they capture the harms of subordinating speech.

At a general level, it's easy to recognize that some speech—hateful, anti-immigrant propaganda, for example—can have oppressive effects, like employment discrimination, or even racist violence. This might be the case when we can trace a causal line from the distribution of hateful ideas about immigrants and racialized groups to discriminatory or violent actions made against members of such groups.¹³

However, this causal story suggests that these utterances are oppressive only in virtue of changing the beliefs and attitudes of some hearers, who then go on to act on these newly formed anti-immigrant attitudes and commit discriminatory and violent acts. In this way, the harm is a 'mentally-mediated' harm, meaning that while speech played a causal role, it is via its impact on the minds of some hearers, namely, their attitudes and beliefs, which later guide their (oppressive) actions.

While surely true, on its own, this picture leaves too much out. It alone can't answer many of the pressing questions raised by subordinating speech. For instance, because of the lag between the hateful utterance and the harmful act, important questions about the scope of responsibility remain unanswered. Such questions include whether responsibility should extend

¹³ Moreover, in both the Nuremberg trials and the International Criminal Tribunal for Rwanda, publishers (such as Julius Streicher) and broadcasters (Ferdinand Nahimana) were found guilty of crimes against humanity and genocide, respectively. See Tirrell (2012, 183–86) for analysis specific to Rwanda.

beyond the perpetrator of the act to include the speaker, and if it does, how much responsibility the speaker should then bear. The simple causal story does not seem to have the resources to address these questions. And this shortcoming leads many theorists to look elsewhere in their accounts of oppressive speech.

Catherine MacKinnon, Rae Langton, Ishani Maitra, and others make the additional claim that some speech *constitutes* subordination—it doesn't just cause it.¹⁴ A quick example can illustrate the distinction at play between causing and constituting in the context of speech. Your boastful assertion that you “can eat the whole thing” might *cause* me to bet you that you cannot, but it's my speech act “I bet you five bucks you can't” that actually *constitutes* me betting you. The difference lies, partly, in what is directly attributable to the speech act itself, and what only occurs with the additional support of some further, perhaps non-linguistic, contingent actions. Some utterances *do* certain things, and this is better captured by describing what they constitute—what they are, as speech acts—than what they cause. Many utterances happen to cause changes in a romantic couple's relationship, and it is not totally incorrect to say that when Sheena first asked Blake out on a date, this eventually caused their marriage. But it is only at the ceremony, when they each say, “I do,” that these speech acts constitute their marrying. In other words, only in so saying “I do” in the proper circumstances, do they perform the act of marrying one another.¹⁵

This distinction between causing and constituting is at the core of J. L. Austin's speech act theory, which offers a useful way of analyzing three distinct aspects of a given utterance: the

¹⁴ Others include Mary-Kate McGowen, Jennifer Hornsby, Jason Stanley, Rebecca Kukla, John Michael Ramsey along with many more who, though they don't all explicitly say so, seem to agree with the claim that some speech constitutes a subordinating act.

¹⁵ This is, in part, because the utterance “I do,” is conventional shorthand for saying: “I hereby marry you,” which is an explicit performative in Austin's jargon, which will be explained in more detail below.

locutionary, illocutionary, and perlocutionary (Austin 1962). On the Austinean model, the *locutionary* content is the (literal) meaning of what's said; the *perlocutionary* effects are what this speech causes to occur; and the *illocutionary* act is what's performed in saying those words in those circumstances.

Consider how, in the wake of the 2017 Québec City Mosque shooting, released documents revealed that the one thing that set the shooter on his murderous path was Canadian Prime Minister Justin Trudeau writing in a tweet: “To those fleeing persecution, terror & war, Canadians will welcome you, regardless of your faith. Diversity is our strength #WelcomeToCanada” (Cecco 2018). By his own admission, this pro-refugee message angered the shooter to the point where he thought it necessary to commit mass murder. One could therefore claim that this tweet was one of the causes of tragic Islamophobic violence and murder, and hence, oppression. And there would be some truth to this claim when we're asking what were the perlocutionary effects of this speech act.

So, while some harmful actions can be caused by speech in the sense that the speech *persuaded* or *influenced* one or more of its hearers to later act in a certain way—in such cases, both the persuasion and the later act would be *perlocutionary* effects of the original speech act. But Langton and others argue that other harmful acts are not so ‘mentally mediated’ and do not fit this simple pattern.¹⁶ Sometimes the utterance itself, as an *illocutionary act*, is an act of oppression, not only because of what effects it eventually brings about, but what the speech does, in and of itself.

¹⁶ And note that this mental mediation need not be so rational as persuasion supposes. For example, repeated speech can also condition hearers to act in certain ways, and this would still fall into a causal model. See Langton (2012).

MacKinnon was perhaps the first to argue that some speech—in her case, pornography, which is legally considered a form of speech in many countries—constitutes oppression in the sense I am alluding to here (MacKinnon 1987).¹⁷ In the philosophy literature, Langton took up this idea, explicitly combined it with Austinean speech act theory, and argued that because pornography can be said to unfairly rank women as inferior, legitimate discriminatory behavior towards women, and unjustly deprive them of important powers, it is therefore speech that *constitutes* subordination (Langton 1993).

To make this argument, Langton uses the example of an apartheid-era South African legislator to demonstrate how *speech itself* can constitute subordination. She asks us to consider the legislator who, in saying “Blacks are no longer permitted to vote,” makes it the case that black South Africans can no longer vote (Langton 1993, 302–03).¹⁸ Her claim is that this utterance—the very act of saying those words in those circumstances—constitutes subordination because it itself successfully *performs* the act of, among other things, unfairly ranking Black South Africans as inferior to white South Africans. That is, beyond merely contributing causally to keeping Black South Africans away from the poles by saying his words, simply *in* saying what he does, the legislator successfully ranks his targets as inferior. Similarly, because these three *illocutionary* acts are performed in the pornographer’s speech acts, pornography is itself an act of oppression of women, and not merely a depiction of it (locution), or an upstream cause of it (perlocution), though it can be those as well.

¹⁷ Althusser (1971) made a similar, though distinct, claim concerning how language is a tool of oppression. Luce Irigaray (1985) is another figure who brought out the oppressive aspects of language forcefully. Pinning down the philosophical origins of the claim that language oppresses is no simple task. Here I aim only to trace the development of a particular version of this claim as it emerged in what might be called feminist philosophy of language, which wed MacKinnon’s claims to Austinean speech act theory.

¹⁸ Langton borrows this example from MacKinnon (1987, 202).

For Langton and others who take up this framework, the difference between causing and constituting lies in the fact that subordination is not just a remote, contingent perlocutionary effect of the speech acts typical of pornography, but is instead an immediate, direct illocutionary act performed in the speech act itself. Pornography doesn't only lead to some downstream acts of oppression (as it has arguably been proven to do), it also *enacts* subordination directly, and is therefore, under the right circumstances, a subordinating act on its own. Similar statements can be (and have been) made for racist or sexist hate speech, which might not only cause subordination, but could also constitute subordinating acts.

1.2. Why Focus on Constituting Instead of Causing?

It is important to note that, in principle, none of the above is meant to undermine the claim that speech can also cause harm in the sense of perlocutionary effects. The claim that speech sometimes constitutes harm may sit comfortably alongside the claim that speech sometimes causes harm. They are complimentary, rather than in competition—at least in theory. As Mary Kate McGowan notes, however: “[a]lthough such causal connections [to harms] no doubt exist, they are very difficult to prove” (2009, 389).¹⁹ As such, the constitutive thesis plays an important evidential role for many theorists, supporting the claim that speech harms even when the causal connections might be in dispute.

Aside from this evidential role, the cause/constitute distinction also matters for two specifically moral reasons, as well, in many accounts. First, as I suggested above, speech that

¹⁹ She adds: “Moreover, even if such causal connections could be established, such oppressive causal effects may simply be the (high) price we pay for freedom of speech. After all, since these (alleged) causal effects are the result of persuasion and hence ‘mentally-mediated’, the speech causing them is highly protected by the First Amendment to the U.S. Constitution” (McGowan 2009, 389).

constitutes subordination can morally implicate the speaker in the harm created by the speech in a more direct way than causal stories, which often leave the speaker one step removed from the harm. Isolating the causal influences of any outcome is always a tough task, especially in the messy world of real life. On some definitions of “cause,” to say a hate speaker is causally responsible for some particular harm is to say that *but for* the speaker’s action, the harm would not have occurred. Yet, for many of the harms these theorists want to attribute to speech, this degree of certitude—or causal effectiveness—is not what’s being asserted. To say a hate speaker’s words themselves constitute a harm, whether or not they’re part of causal chain that results in some specific physical harm, is a distinct claim with different truth conditions.

Secondly, constitutive accounts direct our attention differently than causal accounts. For some speech, the idea goes, what matters is the enactment of certain oppressive norms rather than the discrete injury or injuries they cause. In other words, the target of analysis need not—and perhaps, for strategic or pedagogical reasons, should not—focus mainly on hate speech’s role in bringing about vivid acts of physical violence, but on the normative changes they institute. Therefore the analysis should instead focus on either those subordinating acts that are performed immediately, as in Langton’s example of ranking people as inferior, or, as Maitra puts it, how speech can “constitut[e] norms that help to construct social reality for the subordinated group” (2012, 99).²⁰

²⁰ To further clarify, she immediately adds: “More specifically, these norms determine, first, the (relative) social status of the subordinated group; second, what rights and powers members of the group possess; and third, what counts as acceptable behavior towards those members” (Maitra 2012, 99). In this way, Maitra is looking to echo Langton’s claims about ranking as inferior, depriving of rights and powers, and of legitimating discriminatory behavior. However, because Maitra often implicitly appeals to notions of speech ‘practices’ rather than ‘acts’ as constituting subordination, there does appear to be some distance between her view and Langton’s. I return to this at the end of this chapter.

It is these purported analytical benefits—attributing responsibility to the speaker and fixing our attention on the normative changes in social reality—that help explain the focus on the illocutionary aspect of subordinating speech for many philosophers working on this topic. That is, these aims clarify why *constituting* subordination, rather than simply causing it, is the stated goalpost for many.²¹

However, as I will argue in section 3 below, ambiguities in the concept of illocution weaken its usefulness. I argue that the cause/constitute distinction breaks down, especially for a phenomenon such as subordinating speech, and therefore that illocution cannot do the job that philosophers have sought for it to do here. I show this in part by highlighting the abstract and idealized assumptions underlying the illocutionary approach and argue that a lack of paradigm cases of oppression against which to measure subordinating speech acts poses difficult problems for this model.

I argue, though, that this is no great loss, as philosophers are wrong to put too much emphasis on illocution as the main mechanism of subordinating speech. I claim that it has the drawback of isolating speech acts from their larger context, and in doing so separates the harms of subordinating speech from the harms of oppression more broadly. The illocutionary model does so because it fixates on individual speech acts made by individual speakers. It therefore has limited applicability for the greater terrain of subordinating speech.

This is not to say that illocutionary acts play no role in the analysis of oppressive speech, but rather that it cannot serve as the primary conceptual tool for understanding this problem. I

²¹ In my experience teaching these topics to undergraduate students at US universities, the rationale for the focus on constituting rather than causing subordinating acts is not obvious to the uninitiated reader. This is why I try to make these reasons explicit in the above. This is not to claim, however, that these reasons are the ones that any of the authors whose work I mention would themselves cite.

argue that it, in the end, describes one part of the larger whole that is harmful speech. But a more complete analysis moves beyond the confines of illocutionary acts. We must attend to the harms caused by speech acts as well as those constituted by them. These, along with the other non-linguistic practices of oppression that give subordinating speech acts their meaning, must always be kept in view. This is a difficult task, but a necessary one. As Lynne Tirrell notes: “Attention to language discloses it as a causal force behind actions while revealing how our discursive practices play a constitutive and normative role in what those actions are and in who we are. Language always acts in concert with collateral social practices” (2017, 142). Moving beyond the illocutionary aspects of speech, and explicitly including the broader social practices that function alongside oppressive speech in our analysis is a necessary step in the goal of understanding the harms of subordinating speech.

Furthermore, including both the harms caused and those constituted by speech acts is closer to the experience of the targets of subordinating speech. As Katharine Gelber and Luke McNamara state while summarizing their study on the effects of hate speech as told from the victims’ perspectives: “there is a close and complex relationship between constitutive and consequential harms,” and “attempts to draw a neat distinction between these two types of harm risk misrepresenting, and being insensitive to, lived experiences of public racism” (2016, 336–37).²² To avoid this misrepresentation, my approach throughout this project aims to keep the lived experience of subordinating speech in view. I explore how the tools of philosophy of language—specifically speech act theory—can help explain and describe how speech perpetuates

²² In full, they write: “Our analysis suggests that there is a close and complex relationship between constitutive and consequential harms, and the harms are experienced cumulatively. Attempts to draw a neat distinction between these two types of harm risk misrepresenting, and being insensitive to, lived experiences of public racism. The debate over hate speech laws tends to focus on discrete single incidents and assess the legitimacy of legal sanctions attached to that event. Although that may be necessary to enforce a law, it is an artificial and inaccurate way to understand the experiences of being on the receiving end of hate speech” (Gelber and McNamara 2016, 336–37).

and maintains these structures of oppression. And where the tools need to be modified to fit the account of oppression, I take the facts of oppression as primary and modify the tools, first by rejecting the rigid separation of illocution and perlocution characteristic of many current accounts.

I believe that by starting with the observations of oppression theorists that oppression is manifested most clearly not in individual acts or obstacles, but the systemic and interdependent constraints members of subordinated groups face, we're led to a distinct approach to the harms of subordinating speech than is found in much of the philosophical literature. This approach conceives of subordinating speech as one practice among the larger, interconnected practices of oppression, and this is not something the concept of illocution is capable of capturing. To show this, in the next section I explore the myriad harms of oppressive speech, including the place of illocutionary acts within that larger space. Then, in section 3 I argue that illocutionary acts cannot serve as the primary mechanism for illuminating the harms uncovered in section 2.

2. THREE AREAS OF HARM

To capture the numerous harms of subordinating speech, it is useful to make distinctions among this massive category. For example, in their *Understanding Words that Wound*, Richard Delgado and Jean Stefancic (2004), legal theorists working within critical race theory, note that hate speech can differ along a few distinct axes. These axes include *to whom the speech is directed*: whether an individual, a small group (e.g., a Black fraternity), or at an entire identifiable group (e.g., all Indigenous Peoples). It can also be distinguished in *how it is delivered*: orally, in writing, on the internet, or it can even take a more tangible form, “such as a

monument, flag, or sports logo” (2004, 11).²³ Additionally, they suggest there are a ways that we can distinguish *types of hate speech*: “direct and indirect, veiled or overt, single or repeated, backed by authority and power or not, and accompanied by threat of violence or not” (ibid.).²⁴

In this chapter I distinguish the main areas where the harm occurs.²⁵ I focus on three distinct areas: (1) the psychological health of those targeted by subordinating speech, which often manifests and persists long after any singular instance; (2) the damage inflicted in the performance of the speech act itself, which is identified with the illocutionary act; and (3) the harm done to members of the social group(s) targeted in the subordinating speech *as members of a group*, that is, the hierarchical relations between social groups that partly constitute oppression.

The term ‘area’ is, of course, metaphorical. However, I use it to highlight how each type of harm corresponds roughly to a distinct spatial location, starting with the individual target of subordinating speech in (1); broadening to include both the target, the speaker, and other audience members in (2); and then finally extending outward to encapsulate entire social groups in (3). While there is no doubt some ambiguity in how I draw these boundaries, it is my hope that by first cataloguing these harms separately, we can then better see how the forces that leave individuals and groups vulnerable to these harms in fact connect and mutually sustains one another in key ways.

In addition to cataloguing these harms, another goal of this chapter is to challenge the approach implicit in many leading accounts of subordinating speech, represented below by

²³ See Veilleux (1995) for an analysis specific to sports mascots. For state-erected symbols like monuments, see Tsai (2016).

²⁴ I go on to discuss different types of subordinating speech explicitly in Chapter Three.

²⁵ This is therefore similar to how Delgado (1993) distinguishes the harms of hate speech in terms of (a) their psychological, (b) their sociological, and (c) their political effects. See also Lawrence (1993) who distinguishes the harms of hate speech in terms of (a) the direct psychic injury assaultive speech causes, (b) the reputational libel that can result, and (c) the denial of equal opportunity created by discriminatory environments.

Langton and McGowan. As I interpret these accounts, they assume that a focus on illocutionary acts can help shed light on the structural features of oppression. In section 3, I argue that, in fact, we cannot learn much about the structural features of oppression through the analysis of the illocutionary acts of subordinating speech. That is, the assumption that a focus on illocution can illuminate oppression deserves more scrutiny than it has so far received. And aside from the occasional cursory remarks, the psychological effects of subordinating speech are rarely the focus of philosophical work in this area.²⁶ This neglect, however, is detrimental to the goal of a full understanding of the power of subordinating speech.²⁷ In an effort to remedy both these failures, I first describe the plausible harms of subordinating speech in more detail below.

2.1. Psychological Harms

2.1.1. Immediate harms

Hateful speech has the power to inflict significant psychological harm. Delgado notes that “immediate mental or emotional distress is the most obvious direct harm caused by racial insult” (1993, 93–94). “Without question,” he adds, “mere words, whether racial or otherwise, can cause mental, emotional, or even physical harm to their target, especially if delivered in front of others or by a person in a position of authority” (ibid.). While the particular harms brought

²⁶ The caveat implied by ‘philosophical’ is meant to exclude the groundbreaking work of many leading legal scholars in this area, which is often more focused on the harms caused, rather than constituted by, hate speech. See Matsuda et. al (1993); Delgado and Stefancic (2004); Lederer and Delgado (1995).

²⁷ Nancy Bauer makes a good case that this aim for philosophical precision, whether motivated by humility or something else, has actually turned philosophers away from the crucial questions they need to be asking (and that philosophy is well-suited to try an answer) in this domain. See Bauer (2015, esp. Chapters 5–7).

about by any individual instance will vary from case to case, along with the specific reasons why that harm occur, some generalities can still be uncovered.

One of the most important features to note when discussing the immediate harms of hate speech is how there is often a violent threat that accompanies these utterances—if only an implicit one. Of course, often it is an explicit and direct threat. Other times it is a thinly veiled threat where (historical) context fills in the crucial details, as in the burning of a cross or the hanging of a noose on a Black family’s front lawn.²⁸ It is for this reason some theorists see hate speech *in general*, as a type of threat. Legal theorist Charles Lawrence III, for instance, draws a tight parallel between cross burnings and racist hate speech:

It [cross-burning] is a threat; a threat made in the context of a history of lynchings, beatings, and economic reprisals that made good on earlier threats. [...] The Black student who is subjected to racial epithets, like the Black person on whose lawn the Klan has burned a cross, is threatened and silenced by a credible connection between racist hate speech and racist violence. (1993, 79)

While the above examples are rather direct, it is worth noting how even less overtly threatening utterances like the calls of “Go back to Africa!” or “Muslims stay out!” have at least an implicit threat of further action if the target doesn’t “go back” or “stay out.” I will have more to say about the harmfulness of threats *as an action* in the following section. Here, however, I want to dwell on and emphasize the fact that being the target of a violent threat like this is often a traumatic experience, and thus, a serious psychological harm. Moreover it is a harm that ought to be at the forefront of any analysis of hate speech.

As Lawrence explains, experiencing such hostile speech is like an “attack,” or “a slap in the face.” And furthermore, one which “produces an instinctive, defensive psychological

²⁸ In *Virginia v. Black* the US Supreme Court took up the question of whether a burning cross is in itself a threat, concluding that it is insofar as it is directed at an individual with the intent to intimidate.

reaction. Fear, rage, shock, and flight all interfere with any reasoned response” (1993, 68).²⁹ Moreover, slurs like the N-word function like a “preemptive strike” and tend to “produce physical symptoms that temporarily disable the victim, and perpetrators often use these words with the intention of producing this effect” (1993, 68). In other words, racist, homophobic, and other bigoted epithets are hurled at their targets and, just like the weapons they are, tend to wound those on their sharp end.

This harm, moreover, isn’t simply the visceral result of being unexpectedly attacked—though, importantly, there is that too—but is rationally informed. One crucial source of this harm comes from what Lawrence above calls the “credible connection between racist hate speech and racist violence.” This brings with it an additional, correlated harm, in that this can force the victim to silently listen to the hate speech directed at them. That this is itself a harm might be doubted, but I mention it here because of what it reveals: this being the fact that—despite what many free speech liberals say—*speaking back is a dangerous option*. And taking note of the implicit threat in many hateful utterances helps us see why. As Delgado and Stefancic point out, “many hate crimes have taken place when the victim did just that—spoke back to the aggressor and paid with his or her life” (2004, 13). As such, hate speech throws its targets into this difficult position, where they must very quickly read the situation and choose how (or if) to respond to an unwanted interruption.

²⁹ From this analysis, Lawrence builds towards an account of hate speech that is (or ought to be) unprotected under the first amendment, basing his reasoning on that found in *Brown v. Board of Education*. As he argues: “Psychic injury is no less an injury than being struck in the face, and it often is far more severe. *Brown* speaks directly to the psychic injury inflicted by racist speech in noting that the symbolic message of segregation affected ‘the hearts and minds’ of Negro children ‘in a way unlikely ever to be undone.’ Racial epithets and harassment often cause deep emotional scarring and feelings of anxiety and fear that pervade every aspect of a victim’s life” (Lawrence 1993, 74).

Moreover, even when a target of hate speech deems it worthwhile to speak back, they may find themselves at a loss. As Lawrence notes, “the visceral emotional response to personal attack precludes speech” (1993, 68). It is for this reason that Caroline West claims that racist hate speech “does not function as an invitation to conversation” (2012, 235). And while such a remark may seem trivially true, it does fly in the face of the standard liberal response offered as an antidote to the harms of hate speech—namely, more speech.

Laura Beth Nielsen argues this solution rests on bad empirical assumptions, as most targets of harmful speech make the informed decision to *not* respond.³⁰ As she notes “for the most part, targets of racist speech do not counter it,” thus shouldering the burdens of the interaction on their own (2012, 155). One commonly cited reason is “being fearful for their safety” (ibid., 156).

Even where there is a direct response, Nielsen says “it is hard to imagine that these interactions are changing the mind of the original speaker. After all, hurled epithets are probably not intended to start meaningful dialogue” (ibid., 159).³¹ More plausibly, hate speech of this sort is often intended to inflict harm on its targets. But, regardless of what the speaker intended, from the target’s perspective it is “experienced as a blow, not a proffered idea,” as Lawrence explains, “and once the blow is struck, it is unlikely that dialogue will follow” (1993, 68). It is not merely a threat, but a threat already fulfilled.

Nielsen concludes that this fear of further violence, along with other reasons why people don’t respond to racist speech “demonstrate the difficulties associated with the prescription for

³⁰ Also, it’s worth noting that white men are the most prone to respond to race-related speech, often trying to ‘educate’ the original speaker. See Nielsen (2012).

³¹ One of the more fascinating aspects of Nielsen’s article is how she shows that street begging is actually quite regulated as a form of speech, even standing up to judicial scrutiny in some cases. She says this is mostly likely a result of the fact that, unlike most racist/sexist speech, it “is a request made by the more disadvantaged to the more privileged” (Nielsen 2012, 155).

more speech to counter race-related speech in public” (ibid., 157). But even if one were to dispute her conclusion, that the targets of hate speech often experience fear and other painful emotions could hardly be in doubt. This is distressful, over and above how being threatened can cause instinctive and automatic psychological reactions. “The recipient of hate messages,” Mari Matsuda says, “struggles with inner turmoil” (1993, 24).

Hate speech therefore immediately puts its targets into a multiply painful position, and this is no accident. As West notes, “hate crimes are typically preceded and accompanied by racist hate speech; *and both victims and perpetrators generally know this*” (2012, 234, emphasis added). The connection between hate *speech* and hate *crimes* therefore plays an evidentiary role for its targets, licensing their fear and other forms of psychological distress. “Racist hate speech,” West adds, “is the clearest expression of the kind of hostility that fuels racial discrimination and violence, short of the acts themselves; and it indicates to its targets the presence of these attitudes and behavioural dispositions in the speaker” (ibid.). In light of this, it’s easy to see how forcing individuals to make potentially life-or-death decisions, to assess the severity of the threat, and balance their security against other potential values—like standing up to intimidation, for example—demonstrates the immense psychological toll hate speech exacts in its immediate aftermath.

2.1.2. *Lasting harms*

But even when hate speech is not the potential prelude to a violent hate crime, and is perhaps not even interpreted as a threat, many other psychological and physiological harms remain. “The negative effects of racist hate messages are real,” writes Matsuda (1993, 24).

Victims “experience psychological symptoms and emotional distress ranging from fear in the gut to rapid pulse rate and difficulty in breathing, nightmares, post-traumatic stress disorder, hypertension, psychosis, and suicide” (ibid.). According to studies surveyed by Delgado and Stefancic harms include not only “short-term harms” but also more long-term physiological harms such as “drug-taking, risk-taking behavior, [...] suicide, [...] damaged self-image, lower aspiration level, and depression” (2004, 13).

Taking a closer look at some of the lasting psychological effects of hate speech, Delgado and Stefancic note that these often include: “fear, nightmares, and withdrawal from society” (ibid., 14). For example, the victim of hate speech “may behave circumspectly, avoiding the situations, places, and company where it could happen again” (ibid.).³² And since racism, sexism and other forms of oppression are nearly ubiquitous, this can have a massive impact on the target’s life. Matsuda also notes the far-ranging scope of the consequences of hate speech when she writes:

Victims are restricted in their personal freedom. To avoid receiving hate messages, victims have to quit jobs, forgo education, leave their homes, avoid certain public places, curtail their own exercise of speech rights, and otherwise modify their behavior and demeanor. (1993, 24)

Still other targets respond “with anger, either internalized or acted-out” (Delgado and Stefancic 2004, 14). And others “may reject identification with their own race” or, conversely, “may affect a kind of false bravado” (ibid.).

Children are a special case worthy of extra attention, as they “are among the most easily damaged by racial epithets and name-calling” (ibid., 14–15). West expands on this troubling fact

³² Similarly, West suggests that “racist hate speech may cause those it targets to withdraw from public life and discourse” (2012, 237).

and elaborates on how being the target of hate speech can have negative effects on one's self-esteem, particularly in young children:

If others repeatedly tell you that you are worthless or contemptible—if they say you are dumb dirty, or lazy, simply in virtue of your race which you are powerless to change or conceal—then it is likely that eventually you will come yourself to believe that this is so, especially if the message of inferiority is reinforced in subtle and not so subtle ways by the culture at large.³³ (West 2012, 236)

There is little doubt that stigmatization like this has lasting psychological effects upon its recipients, including “feelings of humiliation, isolation, and self-hatred” (Delgado 1993, 91).

Furthermore, theorists point out how the use of slurs can also trigger more explicit messages of unequal worth, in part because “they draw upon and intensify the effects of the stigmatization, labeling, and disrespectful treatment that the victim has previously undergone,” with the result being “long-term emotional pain” (ibid., 94). In other words, immediate and lasting psychological harms combine with novel utterances functioning to unearth old wounds.³⁴ As

Gelber and McNamara report from their qualitative study with targets of hate speech:

harms are often enduring and not ephemeral. One does not easily ‘shake off’ a racist slur, especially when the encounter and its memory are fresh the next time the target hears of another attack or reads generally circulated hate speech about his/her community. Our analysis suggests that there is a close and complex relationship between constitutive and consequential harms, and the harms are experienced *cumulatively*. (2016)

In sum, an important reason to not skip over these physical and psychological harms too quickly, beyond their significance in and of themselves—especially for the individual victims—is that they contribute to structural inequalities in important ways. That is, these harms “go

³³ Parenthetically, she adds: “One needs only to think of the well-known effects on children of parental verbal abuse to get a sense of the way in which racist hate speech may impact psychologically on its target; and indeed there is evidence that some of the effects of racist abuse on its targets are not dissimilar to the short-term effects of sexual abuse on children” (West 2012, 236). And note that on this latter point, West cites Sumner (2004, 160).

³⁴ As Regina Rini argues in the case of microaggressions, it is precisely because these occur as part of larger patterns of similar utterances that we fail to appreciate their significance if we choose to only examine them as isolated utterances. As she says: “What makes microaggression distinctively harmful is victims’ awareness that each instance is not an isolated accident” (2018, 3).

beyond damage to the psyches and bodies of its victims” (Delgado and Stefancic 2004, 15).

While these harms first emerge in the target’s psyche, they extend outward and may affect the victim’s whole life. And sometimes the full effects of emotional and psychological harms of hate speech are hidden in plain view.

As Delgado and Stefancic remind us, “the person who is timid, bitter, tense, or defensive as a result of frequent encounters with racism [via racist hate speech] is likely to fare poorly in employment and other settings, as well” (ibid.). The impact of stereotype threat is relevant here as well, as encounters with hate speech tend to make one’s group identity—and the pernicious negative stereotypes many attach to that group identity—salient where it perhaps was not before.³⁵

But seeing these harms as the possible or probable effects of certain instances isn’t to say that they are the result of every instance of hate speech, even instances that are explicitly hostile and backed by force. The targets may, in some cases, have particularly thick skin, and just aren’t the sort of person to be intimidated by a threat. And there is likely something highly specific about how any single person responds to different cases of harmful speech. Individuals are differentially sensitive not just in general—thick vs. thin skinned—but to particular instances.

Moreover, these causal, psychological harms of hate speech are most vivid when performed in face-to-face interactions and might be dissipated partially when performed via different mediums. But where these harms might not occur in every instance or may dissolve in different contexts, other harms remain.

³⁵ For more on stereotype threat, see Cordelia Fine’s (2010) *Delusions of Gender*.

2.2. Harmful Illocutionary Acts

The preceding provided a short overview of some of the common psychological harms that can be the causal effect of being the target of subordinating speech. Legal theorists—specifically critical race theorists—have paid the most attention to these harms, in part because of their aim to make legal cases for actionable torts and the prohibition of certain forms of hate speech. These harms are not so much neglected by philosophers writing on subordinating speech as they are underemphasized and underexplored. Overall, philosophers pay greater attention to harms that are inflicted in the *performance* of the speech act itself—those that comprise the illocutionary act. McGowan makes this shift in focus away from causal harms explicit when describing her approach to oppressive speech: “rather than focus on what a certain category of speech *causes*,” she writes, “I am interested in what such speech actually *does, in and of itself*” (2009, 389–90).³⁶

Whereas the causal harms can be thought to follow from the speech acts, the illocutionary harms are claimed to be the more direct, less contingent result of the utterance. That is, while it is certainly true that *by* saying “Jews will not replace us!” a pack of marching white supremacists can cause the fear, heightened anxiety, and raised blood pressure of their Jewish (and non-Jewish) onlookers, it is a further and distinct fact that *in* saying these words the speakers do many other things.³⁷ Some conceivable examples include: they *rank* Jewish people as inferior; they *discriminate* on the basis of ethnic/religious identity; and they might even *subordinate*

³⁶ Similarly, Langton notes the following when discussing the central focus of the anti-pornography ordinance that partly inspired her work: “It was about a harm enacted by pornography as a speech act, an ‘illocutionary’ act, in terms introduced by J.L. Austin— a harm distinct from its content, as ‘locutionary’ act, and its effects, as ‘perlocutionary’ act. The claim was that pornography can enact harm— just as an oppressive law can enact harm” (2017, 24).

³⁷ The distinction between acts performed *by* saying X and those performed *in* saying X is first laid out, as a helpful heuristic device, by Austin (1962).

through their words. Plausibly, through their utterance they also *silence, terrorize, demean, debase, debilitate, degrade, assault*, and so on. Beyond that, other, similar utterances might *ridicule, humiliate, exclude, outgroup, disrespect*, and much more.³⁸

All the above verbs are meant to identify different illocutionary acts that are enacted *in* saying certain words in particular circumstances. They identify speech acts that, even though they often imply future downstream consequences, are not dependent on harmful future effects for being acts of harm themselves. Therefore, the illocutionary harm is identified with what the speech act constitutes, rather than what it causes, recalling the distinction made earlier.³⁹ In other words, when the utterance itself constitutes a harmful act, it is a harm, whether or not it also causes further *physical or psychological* harms.

Take, for instance, threats. Above I examined some of the common psychological effects of being repeatedly targeted by racist hate speech, which some consider threats. And yet, each individual instance of a threat might not result in any of these harms. Perhaps the target might have so-called thick skin, or perhaps the details of the situation render the speaker less imposing than they might otherwise be. A lone, elderly white man targeting his hate speech at a large group of younger gay men might not succeed in making his audience feel threatened at all, and thus they likely won't incur the psychological costs described above.

There is nonetheless a difference between *feeling* threatened, and being threatened—or, more precisely, between being threatened and experiencing the usual negative psychological effects of being threatened. And just because an individual might not experience any specific

³⁸ I go into greater detail about the sheer variety of subordinating speech acts, including where they overlap and reinforce one another, in Chapter Three.

³⁹ This is so even while most writers on this topic acknowledge some inevitable overlap and ambiguity between what is caused and what is constituted. See, for example, McGowan (2009, 404–05). I go on to explore this tension in the following sections. At the moment I aim only to explain the central claims being made.

psychological harm in a given instance doesn't mean we should conclude no harmful act was performed. To threaten someone is an illocutionary act, and it may be harmful irrespective of the possible causal harms discussed above. Like promises, threats fit into the category of illocutions that Austin labels "commissives." These are speech acts that "commit the speaker to a certain course of action" (1962, 150–51). But Austin makes clear that their success as speech acts is independent from their success at predicting the future, like whether the speaker follows through with their promise or threat.⁴⁰ Much like how it's not necessary for me to actually carry out your commands for you to successfully order me, it's not necessary for you to later assault me for you to successfully threaten me.

This isn't to say that every instance of a particular utterance constitutes a threat and therefore a harm. The context, or total speech situation, matters for identifying illocutionary acts, and plays a role in whether an utterance is a threat, and whether that threat is harmful. Even saying "I'm threatening you right now" wouldn't always amount to a threat if it occurred during a play, for example. But like most speech acts, threats are not fixed *solely* by the audience's reaction. This is not because those effects do not matter; far from it. Rather it is simply that they are not *all* that matters. The shift in normative landscape also matters, and this occurs largely via the utterance itself, and not solely because of any later effects.

In what follows, I focus on the illocutionary acts of 'enacting,' 'inciting,' and 'ranking.' Enacting because it plays a particularly important role in McGowan's account of subordinating speech, inciting as it has been the focus of many legal prohibitions on hate speech, and ranking because that is perhaps the fundamental speech act of subordination on this illocutionary

⁴⁰ This elides over a bit too quickly the important role that 'uptake' plays in fixing, even partly, what speech act was performed. For example, if I warn you to watch out for bears, but you think I'm joking and therefore don't take it as a warning, was my utterance, in fact, a warning? I return to this issue below.

approach. Moreover, they each illustrate another of Austin's categories of illocutions, in addition to 'commissives' discussed above; namely, 'exercitives' and 'verdictives.'⁴¹

2.2.1. *Exercitives: Enacting oppressive rules*

In a series of articles, McGowan argues that speech can itself "enact" discriminatory rules, and hence oppression. Moreover, she claims that speech act theory helpfully reveals this illocutionary aspect of (some) racist and sexist utterances (McGowan 2003; 2004; 2009; 2012). Her account explores and expands the type of speech act Austin named "exercitives." The defining feature of an exercitive is that it "enacts permissibility facts" (McGowan 2009, 392)—which we might also call 'rules' or 'norms.'

Examples of exercitives are easy to come by, especially in relationships where some form of hierarchy is present. A parent, for example, telling their child "no more games before you finish your homework," enacts the fact that it is now impermissible for the child to play any further games until they finish their homework. When the boss announces that "any employee who leaves before the store is clean will be fired" enacts the new permissibility facts that it is impermissible for the employees to leave early, as well as the fact that it is now permissible for the boss to fire anyone who does.⁴²

⁴¹ Along with those mentioned above, 'silencing' has also received special attention in the literature, particularly within feminist literature about the harms of pornography. This literature—too vast to summarize here—begins with MacKinnon's observation that there are "words that set conditions" for other speech act's success (1993, 63–68). That is, there are "some kinds of speech [that] can set the conditions for other kinds of speech: they make some speech acts possible for some, and impossible for others" (Hornsby and Langton 1998, 27). Pornography, they go on to argue, does just this, and inhibits the speech of women. That is to say, "pornographic speech acts help create a communicative climate in which the felicity conditions for some of women's speech are not met" (ibid.).

⁴² It is important to note that these are not (necessarily) legal or moral permissibility facts, but rather more local and context-dependent facts.

McGowan adds that if the rules that are newly enacted through an exercitive are themselves discriminatory rules, then that speech act *constitutes* an act of discrimination. And if it is discrimination on the basis of race, then utterance is an act of racial discrimination. As such, when a Black man steps onto a bus filled with only white people, and a white man utters: “Turn around boy, you’re not welcome here,” that man enacts the (local) rule that the Black man is not permitted on the bus—or at least he tries to.⁴³

What some racist speech does, then, is functionally identical to what a “whites only” sign hung up in a business does, making it impermissible for people of color to get service at this particular business.⁴⁴ And since the latter is an illegal act of racial discrimination, so is—or at least might be—the former (McGowan 2012, 136–38). And since racial discrimination is one aspect of broader systems of racist oppression, McGowan concludes that “speech oppresses in virtue of enacting permissibility facts (that oppress)” (2009, 392).⁴⁵

In Chapter Three, I will argue that this approach mischaracterizes the majority of subordinating speech. Here, I only want to emphasize how it is important for McGowan that the harm she identifies is not caused by speech but is instead constituted by it. It matters because McGowan says that looking to what some speech *is*—what harmful act it constitutes—triggers a different type of evaluation than assessing the potential harms it might cause. In the latter case, the question becomes whether the benefits of permitting such speech might outweigh the harms that it causes. Here we weigh the pros and cons of regulation. For example, in many nations the

⁴³ Adapted from McGowan (2012).

⁴⁴ It is worth restating that these permissibility facts need not be moral or legal permissibility facts. They can be more local. In this case, what is newly impermissible might be for the employees to serve people of color without sanction from the owner.

⁴⁵ It is worth noting how McGowan puts things in even stronger terms, singling out the exercitive as the illocution of all oppressive speech acts. As she puts it, “[s]ince oppressive speech enacts permissibility facts (that oppress), speech must be exercitive in order for it to be an act of oppression” (2008, 392). In Chapter Three I argue against this view of subordinating speech, when I go into McGowan’s view in more detail.

harms of defamatory speech—as well as the harms incurred by regulating it—are thought to outweigh the benefits of permitting this speech, and so the legal prohibition of defamation is thought justified. But things might be different—or so some claim, at least—for something like racist hate speech, where the cost-benefit analysis may go the other way.⁴⁶ A focus on causal harms leads to the issue of balancing harms and benefits, and this, McGowan argues, leads to an impasse in the debate over hate speech (2009, 389; 2012, 122).

She aims to avoid this impasse by focussing on what harmful speech constitutes. Classifying speech according to what it constitutes, she claims, enables potential legal regulations that would be more difficult within a strictly causal account. Thus, McGowan believes that examining the specifically illocutionary aspect of subordinating speech reveals harms—like racial discrimination—that open up new avenues of understanding and is crucial in developing strategies for how to confront and counter these harms.

Below I will present reasons for skepticism about the viability of this approach. As I see it, the attempt to draw such a sharp line between the illocutionary and perlocutionary aspects of subordinating speech acts is both conceptually impossible and strategically unwise. We cannot carve off the effects of racist speech and hope to focus solely on the act they constitute. The two are inextricably linked. I believe this becomes apparent when we take a closer look at the claim that some speech constitutes harmful acts and put this alongside descriptions of the perlocutionary effects of subordinating speech. I continue to explore these ideas throughout section 2, before turning to criticism in section 3.

⁴⁶ One example is Andrew Altman (2012), who argues that the laws prohibiting Holocaust denial and other anti-Semitic language sweep too broadly, potentially including some of Kant's racist writing. And, he argues, this shows that these laws are currently unjustified.

2.2.2. *Incitement as an illocutionary act*

The idea of classifying speech according to what act it constitutes in effort to potentially regulate certain *types* of speech is not new. In *On Liberty*, John Stuart Mill argued that “even speech” ought to be constrained when its utterance constitutes “a positive instigation to some mischievous act.” They go on:

An opinion that corn-dealers are starvers of the poor, or that private property is robbery, ought to be unmolested when simply circulated through the press, but may justly incur punishment when delivered orally to an excited mob assembled before the house of a corn-dealer, or when handed about among the same mob in the form of a placard. (1859, Chapter Three, paragraph 1)

The claim here is that the type of act being performed when making an utterance can depend on its context. The very same content—that corn-dealers are starvers of the poor—is in one case a mere expression of an opinion, and in another an incitement to violence. In noting this difference, Mill are acknowledging the role of context in fixing what *act is being performed* by an utterance. In each case, the very same words are uttered, but very different speech acts are performed. And crucially, the latter “may justly incur punishment” because of what act it is, while the former ought to always be immune from such punishment—at least by their lights.⁴⁷

Daniel Jacobson makes this point explicit when, in discussing how speech act theory can assist Mill in making this distinction, he notes that “the conception of incitement as an illocutionary act [may] help provide the subtler understanding necessary to distinguish speech that is probably harmful from that which positively incites” (1995, 71–72). That is, while some statements—for example, that private property is theft, or that there is no god—may still eventually lead to harm and violence when merely circulated through the press, it is only when

⁴⁷ A similar point lies at the heart of the 'fighting words' doctrine in the limits of US free speech laws.

such statements themselves constitute an act of incitement that prohibition or punishment is legitimate.⁴⁸

And, like Jacobson does for Mill, Langton suggests speech act theory can help make sense of the underlying thought in the United Nation's prohibition embedded in its *Universal Declaration of Human Rights*. Article 7 states that everyone is "entitled to equal protection against any discrimination in violation of this Declaration and *against any incitement to such discrimination*" (Universal Declaration, emphasis added).⁴⁹ And while here it is incitement to discrimination—rather than violence—that is at issue, Langton claims it is nonetheless incitement as an illocutionary act that is positioned as deserving of special attention and possible prohibition.⁵⁰

The intended target of such a prohibition include historical examples like the Nazi propagandist Julius Streicher's publication *Der Stürmer*, which contained anti-Semitic caricatures along with hateful calls to genocide, or anti-Tutsi broadcasts in the lead up to the Rwandan genocide, which likened the Tutsis to snakes and cockroaches.⁵¹ More proximate examples include the rise of anti-Muslim messaging in post-9/11 US⁵²—up to and including the

⁴⁸ It is worth noting that Hornsby and Langton disagree with this assessment of Mill and go on to claim that the incitement Mill has in mind here is properly distinguished as a perlocutionary effect of certain speech (1998). Nevertheless, in later work, Langton highlights 'incitement' as one harmful illocutionary aspect of hate propaganda (2012), and so she agrees with my broader point, that incitement is, at least sometimes, an illocutionary act.

⁴⁹ For more on the illocutionary dimension of the Universal Declaration's prohibition on certain speech, see Altman (2012), and for a similar analysis of the UN's *International Convention on the Elimination of All Forms of Racial Discrimination*, see Langton (2012).

⁵⁰ That such a prohibition on incitement can sit comfortably alongside a robust freedom of opinion and expression—at least according to the Universal Declaration's Article 19—suggests that it is the *illocutionary act*, and not the *locutionary* content, that is under suspicion. Holocaust denial is, perhaps, a special case, however. As this is one 'opinion' whose mere expression has been deemed worthy of criminal prohibition in many states. And yet, it is worth noting that the discussion of the beliefs and claims of Holocaust deniers is not banned. And this suggest that it is still nonetheless a type of illocution—assertion—that is claimed to amount to incitement, rather than say, reporting. See Waldron (2012) for the claim that, in most cases, the content of the utterances that he believes ought to be prohibited could nonetheless still be expressed, though with less invective language.

⁵¹ See Tirrell (2012) for analysis of the role of language in the Rwandan genocide.

⁵² Waldron (2012) contains many examples and discussion of anti-Muslim hate speech in contemporary US society.

present—and many of Donald Trump’s now-infamous remarks concerning Mexican and South American immigrants.⁵³

The idea behind such special concern for these types of utterances is that speech that constitutes incitement to violence and/or discrimination—especially racial or religious violence and discrimination—is outside the bounds of what a just society should tolerate. It is beyond the pale. And it is beyond the pale because of what it *is* and what it *does*. As Langton writes:

It [racist hate speech] promotes racial hatred and discrimination—‘promotes’ in a causal sense. It also incites racial discrimination and hatred, and promotes racial hatred and discrimination—‘promotes’ in an advocacy sense.... So hate speech ‘promotes’ hatred in both illocutionary and perlocutionary ways: it advocates and causes hatred. (2012, 75–76)

While Langton’s conception of hate speech has both a constitutive, as well as causal dimension, it is primarily the constitutive—illocutionary—dimension that Langton addresses in her analysis. On this analysis, attention to the illocutionary act embodied by these utterances reveals how they perform the act of *rendering a verdict* of inferiority and/or suspicion about some group(s), and *calling for violence* based on retribution, revenge, or deluded sense of self-protection.⁵⁴ Seen in this way, the speech act is what Austin labels a ‘verdictive,’ that is, the act of rendering a judgment. And is it this feature of the speech that Langton suggest renders it harmful and open to regulation.

It is worth pausing to note that here the focus is mainly on hate speech as it functions as propaganda, rather than hate speech as it functions as an assault.⁵⁵ In this capacity, its ‘targets’

⁵³ In announcing his candidacy, Trump said, “When Mexico sends its people, they’re not sending their best. They’re not sending you. They’re not sending you. [sic] They’re sending people that have lots of problems, and they’re bringing those problems with us [sic]. They’re bringing drugs. They’re bringing crime. They’re rapists. And some, I assume, are good people.” See Washington Post (2016) for the full remarks. Saul (2018a) discusses some of the hateful rhetoric of Trump’s campaign.

⁵⁴ The importance of the notion of “white genocide” for white supremacist groups highlights this last feature well, in my opinion.

⁵⁵ Langton (2012) acknowledges this is her focus and refers her readers to the work of Matsuda (1993) for an account of hate speech as assault. This represents the tendency I alluded to above, where philosophers tend to

are not only the racial minorities whom it disparages, but also members of the same racial/ethnic group as the speaker. Indeed, this might be considered its primary audience. Much like how the corn-dealers are not themselves the main intended audience in Mill's example, this type of racist hate propaganda is perhaps best understood as ingroup speech. In this capacity, it looks to spread its message to like-minded and potentially sympathetic ingroup members. This sort of incitement may thus be seen as acts of outreach and advocacy. Jeremy Waldron suggests the implicit message of this type of ingroup hate speech consist in something like the following:

We know some of you agree that these people are not wanted here. We know that some of you feel that they are dirty (or dangerous or criminal or terrorist). Know now that you are not alone. Whatever the government says, there are enough of us around to make sure these people are not welcome. There are enough of us around to draw attention to what these people are really like. Talk to your neighbors, talk to your customers. And above all, don't let any more of them in. (2012, 2–3)

This clarifies the underlying idea of speech as an act of incitement, spelled out explicitly. It both passes judgment and calls on others to action. While these two components of the speech act are related in a direct way to downstream acts they bring forth—whether this is exclusion, discrimination, violence, or even genocide—Langton stresses how this emerges from the act being performed immediately. “The effects are there,” she says, “because of what hate speech is, as an illocutionary act: it incites hatred” (2012, 76). This is implicit in her analysis of the South African legislator's speech, considered above. As she puts it, it is because the legislator's speech *constitutes* an illocutionary act of ranking and legitimating that Black South Africans later suffer further discriminatory abuses. But rather than mere cause and effect, she urges us to see the effects as the result of the enactment of a status, something performed in the illocutionary act.

acknowledge the causal harms of hate speech, but don't themselves engage in much analysis of this aspect of subordinating speech. Chapter Two takes up the speech act approach to 'assaultive' hate speech, which similarly centers on verdictives.

The above analysis provides *part* of the story of the harm found in the illocutionary act. And while it is helpful at first to consider how ingroup members are the primary audience for this type of hate propaganda, it is crucial to examine how the outgroups members receive this speech. I go on to do so now as I continue to explore the illocutionary dimension of the harm of subordinating speech, all with an eye to later demonstrate the limitations of this approach.

2.2.3. *Verdictives: Ranking as inferior*

Though they aren't the immediate audience of the type of ingroup hateful speech considered above, people of color, immigrants, women, etc., are not simply on the sidelines, unaffected by this speech until it spills over into the realm of 'action.'⁵⁶ In addition to the potential incitement of ingroup members, hate propaganda also often has an immediate impact on its indirect audience—e.g., the outgroup members it demonizes and degrades.

To say that speech incites hatred is, in part, to say that it directly contributes to a climate where racist hatred is more pervasive and vibrant. It does so even if no one in its intended audience acts any differently as a result,⁵⁷ but through how it alters “the permanent visible fabric of society” (Waldron 2014, 3). And, according to advocates of the illocutionary model, one way to understand how it accomplishes this is by seeing these speech acts as *verdictives*, whose force is that of an authoritative judgment.

The description of these utterances as “authoritative judgments” might arouse suspicion where the speakers are not themselves in any recognizable position of authority—and so, unlike

⁵⁶ One salutary aspect of speech act theory is how it breaks down speech-action binaries. This is one reason I adopt the terms of speech act theory while also criticizing some conceptions of them.

⁵⁷ Though, it is worth noting, that this is unlikely as an empirical assessment of the effects of hate propaganda.

the legislator case. In these cases, the thought might be: aren't these speakers just hapless bigots, whose own low-status renders their speech more embarrassing than menacing? To be sure, verdictives do standardly require the speaker to occupy some position of authority. To call the runner 'safe,' the speaker must be an umpire, and shouts from the crowd have no impact on the score of the game.

This objection to the force of subordinating speech is called the Authority Problem, and I take it up in more detail in Chapter Two. Here I simply want to note how it would be a mistake to think 'ordinary' hate speakers have no impact on the lives of the people their utterances target, even if it is 'low speech,' more like shouts from the crowd than the umpire's official judgment. Unlike in baseball, where the game operates separately from the crowd, when you live among the crowd—just as marginalized people live alongside the dominant—the words and judgments of these speakers do have real impact.

We see this best when we consider how members of the group(s) targeted by such speech might take up these speech acts. Even if they are not its direct intended audience, they are still likely to encounter these messages—indeed, it's likely that propagandists count on this. First, there is the expressed message plausibly received by the members of minority groups attacked by such speech, which Waldron summarizes as:

Don't be fooled into thinking you are welcome here. The society around you may seem hospitable and nondiscriminatory, but the truth is that you are not wanted, and you and your families will be shunned, excluded, beaten, and driven out, whenever we can get away with it. We may have to keep a low profile right now. But don't get too comfortable. Remember what has happened to you and your kind in the past. Be afraid. (2012, 2)

If left unaddressed by the wider society, according to Waldron, the overall effect of these utterances is the disintegration of whatever formal and informal assurances of inclusion the wider (putatively egalitarian) community endorses. Without these assurances, minority

communities cannot exist as equals alongside the other members of society. And, for Waldron, this corrosive, undermining aspect of hate speech is what reasonably grounds legal prohibitions on such speech for nations committed to the equal dignity of all its members (Waldron 2012, esp. Chapters Three and Four). And for speech act theorists like Langton, it is in understanding hate propaganda as an illocutionary act—one that advocates, promotes, and incites hatred—that reveals how speech acts like these function to *immediately and directly* marginalize members of the targeted groups. This is what these utterances do, constitutively. They are a *direct* attack on the dignity of the members of oppressed groups.

Langton clarifies this issue concerning the illocution of *legitimizing*, also a verdictive act, not too dissimilar from incitement. She acknowledges that “one effect of legitimating something is that people believe it is legitimate.” But, she adds, “they believe it is legitimate because it has been legitimated, not vice versa” (1993, 303). That is, their beliefs are to be explained by the fact that certain behavior has “indeed been made legitimate in that particular arena of activity” (ibid.).⁵⁸ This is so, because “the illocutionary act of legitimating something is to be distinguished from the perlocutionary act of making people believe that something is legitimate” (ibid.).⁵⁹

⁵⁸ Those last few words are crucial, as this is a claim relative to a particular domain, one that itself may not ultimately be legitimate. As she adds in the immediately following, “there may still be some perspective outside that arena from which one can say that discriminatory behavior is never truly legitimate” (Langton 1993, 303).

⁵⁹ Langton further clarifies her views on this in later work by appealing to how she understands the difference between ‘score’ and ‘common ground,’ when she writes: “What Lewis sees as two ways of understanding the score, we can see as complementary. We need both. A change in attitudes is a downstream effect of what is done with a ball, or with words, in the ‘game’ itself. Aiming to fulfill certain regulative rules (about winning the game), the batter scores a home run, as defined by certain constitutive rules (about what counts as a home run). As Lewis puts it, the score ‘straightway’ changes— which means at that very time. The attitudes ‘in many heads’ change thereafter, as a causal consequence. The home run is comparable to an illocutionary speech act; the attitudes to some especially salient perlocutionary acts. Score tracks illocutionary acts (inter alia). Common ground tracks (some) perlocutionary acts. To identify normative score with psychological common ground would be to risk psychologism.” (Langton 2018a).

According to advocates of this model, then, it is by looking at the illocutionary aspect of this speech that best reveals a distinct and important harm of hate speech. This contrasts with attention to either its locutionary content, which doesn't tell us enough about how these words are being put to use, or its perlocutionary effects, which is not specific enough, as it plausibly includes not only increased racial hatred in some audience members, but also indignation and outrage in others. For Langton and McGowan, concentrating on the illocutionary dimension helps to both narrow the focus and draw attention to what matters most in evaluation these speech acts. And to be sure, their work does indeed uncover significant dimensions of the harm of subordinating speech.

As this approach rightly reveals, it's a mistake to consider hate speech solely in terms of offering a disembodied message that competes with legal assurances of inclusivity that some audience members come to believe, and others reject.⁶⁰ To do so uncritically neglects the roles that context, power, and authority all contribute to the harms of hate speech. I explore the topic of speaker authority more fully in the next chapter. But even when bracketing the impact of speaker authority on the harms of subordinating speech there is no shortage of harms attributable to hate speech as an illocutionary act. These are utterances of exclusion. They defame their targets as criminals and/or terrorists,⁶¹ exploiting the ambiguity of generics.⁶²

⁶⁰ And this assumes these assurances even exist; often they do not. Examples of the lack of even formal, legal assurances of equal status abound. From laws against same-sex marriage, to failures to recognize the existence of transgender people, as well as the brunt faced by undocumented persons who are literally smeared with the term 'illegals.'

⁶¹ For Waldron, understanding hate speech as a type of group-defamation or libel provides compelling reasons to justify legal prohibitions on such speech. See Waldron (2012).

⁶² In recent work, Langton offers a helpful explanation of the power of hate speech (and other forms of speech) to harm even without authority. She writes: "Hate speech might wound without authority, because other factors contribute to its force: the oppressive history of slurs, and the power of words to invoke and retrieve that history; the 'truthiness' of danger and conspiracy claims, got from a seeming-truth of the content, not the seeming-trustworthiness of the speaker; the attention-grabbing, evidence-resisting power of lurid rumours, whatever their source; the word-transcending power of graphic images to shape attitudes; the ambiguity of unquantified generics, reifying danger or inferiority, tarring too many with a too-sweeping brush. Hate speech may work to alter perception

Yet, as I will argue below, in drawing such a sharp contrast between illocution and perlocution—and thus, between causing and constituting harm—Langton and others commit two related errors. First, they tie the success of their accounts to the plausibility of such a sharp distinction, which as I will argue is unfounded. Second, and more significantly in my mind, they restrict the analysis of subordinating speech in such a way that it fails to do justice to the actual phenomena under investigation. The first step I make towards making this argument more plausible is to describe my third area of harm—those that occur at the structural level.

2.3. *Unjust Structural Relations*

There is another area of harm that is not captured in the above two. Beyond the psychological harm that sees hate speech as akin to direct assault, and in addition to the illocutionary harms that impact the targets of subordinating speech even when these causal harms are absent, there is the harm done to the broader society. That is, there is a harm done to the members of the *entire groups* targeted by subordinating speech, rather than just those members who happen to be (part of) the speech's audience.

This third area of harm is not fully distinct from the other two, since it serves as a precondition for the precise types of harms the other two areas aim to describe. That is, the psychological harms of racist hate speech are importantly distinct from the psychological harms that may result from similar speech, such as non-racist verbal bullying, which may overlap in many ways with racist speech.⁶³ The psychological harms caused by hate speech, such as

itself, so that we literally come to see our fellows as dehumanized or animal-like, literally hear them as shifty, contemptible, or dangerous. Unlike the law, the harm in hate speech dwells partly in factors independent of its authority” (Langton 2018a).

⁶³ See Kumashiro (2018) for an analysis of when these overlap.

feelings of fear and vulnerability, have unique features when what one (reasonably) fears is assault from a racist society, rather than from specific, known individuals, as is typical in bullying. Moreover, since McGowan's claim is that some speech constitutes the enactment of *oppressive* discriminatory permissibility facts, she takes it as a given that for this to be so, “an unjust social arrangement is nevertheless required” (2009, 391). In this way, both the (causal) psychological harms and the (constitutive) illocutionary harms discussed above implicate wider systems of oppression. In other words, they each get part of their force *as harms* from background injustice, and so a complete understanding of either would require an analysis of oppression (McGowan makes this point explicit, see 2009, 398ff).

This section offers a brief, incomplete sketch of such an account—a complete account is, of course, well beyond the scope of this project. In addition to this sketch, the other main goal of this section is to fill in further details about how subordinating speech contributes to unjust hierarchical relations between social groups—that is, how speech *contributes* to oppression.

2.3.1. *Oppression and oppressive speech*

While the literature on the features, mechanisms, contours, and experiences of oppression is expansive, here I intend only to draw out two common themes. First is the fact that oppression affects individuals as members of identifiable social groups. As Marilyn Frye remarks in her classic essay on the topic: “One is marked for application of oppressive pressures by one’s membership in some group or category. Much of one’s suffering and frustration befalls one partly or largely because one is a member of that category” (1983, 15–16). This is evident in the examples considered above. Racist hate speech impacts members of marginalized racial groups

as members of those groups. It is virtue of a perceived identity that hate speakers select their targets.⁶⁴

Similarly, the importance of group membership is evident in the cases of propaganda described above. It is obvious how Nazi propaganda targeted Jewish people *as* Jewish people and how anti-Tutsi propaganda singled out Tutsis as deserving denigration. Acknowledging this amounts to a truism about these forms of dehumanizing propaganda: it targets a group or several groups in order to increase animosity towards those groups. Without a social group to disparage, it would be a different form of propaganda altogether—if we even consider this type of speech propaganda anymore.⁶⁵ The importance of group membership, then, and in particular the visible markers of group membership, must be kept in view in any coherent account of the harms of oppression, including oppressive speech.

The other common element of analyses of oppression I wish to bring out is its structural nature. Again, it is impossible to do justice to the richness and variety of the existing literature. For my purposes, though, what is crucial to bring out is the necessity to conceptualize oppression at the macro-level, in addition to detailing its micro-level harms. While individual instances of violent hate crimes, flagrant sexist discrimination, or angry racist tirades are micro-level instantiation of oppression—its most visible, blunt instruments—it is in virtue of how these micro-level harms and constraints systematically connect into “networks of forces and barriers”

⁶⁴ I say “perceived” here because hate speakers occasionally read their target’s identity incorrectly—mistakenly shouting anti-Muslim content at Sikhs is a common case in North America—and yet it is still true that they targeted that individual in virtue of group-membership.

⁶⁵ At this point, I am agnostic as to whether propaganda that does not serve to disparage any group is still propaganda, or whether it is something else entirely. Environmentalist propaganda might be an example here, but I do not wade into this issue. My reasons are mainly to sidestep thorny questions about the proper definition of propaganda in order to focus on the uncontroversial cases. Moreover, since my aim in this chapter is to describe the varied harms of subordinating speech, I believe it is reasonable to put aside propaganda that plausibly does not harm (as it does not target) any social group.

that oppression becomes the pervasive an all-encompassing injustice that it is (Frye 1983, 3).⁶⁶

What's more, it is important to recognize that there is not one system of oppression—say, class oppression, as the caricature of Marxism goes—that lies at the foundation of all experiences of oppression, but rather there are many, which combine and interlock in myriad of ways. As

Patricia Hill Collins et al explain:

the notion of interlocking oppressions refers to the macro-level connections linking systems of oppression such as race, class, and gender. This is the model describing the social structures that create social positions. [Furthermore,] the notion of intersectionality describes micro-level processes—namely, how each individual and group occupies a social position within interlocking structures of oppression described by the metaphor of intersectionality. Together they shape oppression. (Collins et al. 2002, 82)

Iris Marion Young makes a similar, though distinct point when she offers a “plural explication of the concept of oppression” in *Justice and the Politics of Difference* (2011). As she argues, oppression is not capturable by any set of necessary and sufficient conditions. That is, there is no universal experience of oppression that properly picks out all genuinely oppressed people. The oppressive forces that shape the lives of Black people in the US, Palestinians in Gaza, Indigenous Peoples worldwide, and many more oppressed groups are importantly different, though still related. For Young, a social group is oppressed if its members routinely experience any one of what she calls the “five faces of oppression”—exploitation, marginalization, powerlessness, cultural imperialism, and violence.

In identifying these five systemic practices as being constitutive of oppression, Young pulls the focus away from individual, hateful events—like lynchings—as being the hallmarks of oppression. While it is no doubt true that many instances of oppressive acts are the direct result of someone acting from a hateful motive, Young aims to draw attention to the pervasive and

⁶⁶ Frye's famous birdcage metaphor is a useful illustration of the interrelation between these two levels. For more contemporary accounts, see Cudd (2006); Haslanger (2012; 2015); and Zheng (2018).

everyday oppressive practices that have their causes in “unquestioned norms, habits, and symbols, in the assumptions underlying institutional rules and the collective consequences of following those rules” (2011, 41).⁶⁷ Immense social problems like systemic economic exploitation, pervasive homelessness, and the marginalization of many peoples’ cultures are often the cumulative effects of countless smaller decisions and actions. According to Young, we see these best when our eye is on the structural patterns that are the result of seemingly innocuous transactions.

My purpose in highlighting the two above features of oppression—the group condition, and the structural condition—is to bring out something that I believe is elided over in much of the subordinating speech literature. This is the pair of facts that, first, oppression manifests in myriad and varied ways and therefore is irreducible to any set of core features that could adequately capture every instance, so we are better off looking toward the structural constraints. And second, that the individuals targeted for subordinating speech are harmed *as members of oppressed groups*, and this has important consequences for all other members of these groups.

Without disputing the importance of the harms brought out by the previous two categories already explored, my claim is that many further relevant harms like stereotyping, objectification, Young’s cultural imperialism, and more are not adequately captured by a focus on either the direct psychological harms of hate speech, or its illocutionary force.

⁶⁷ As Young defines it: “structural oppression refers to the vast and deep injustice some groups suffer as a consequence of often unconscious assumptions and reactions of well-meaning people in ordinary interactions, media and cultural stereotypes, and structural features of bureaucratic hierarchies and market mechanisms—in short, the normal procedures of everyday life” (41).

2.3.2. *Social inequality and language: An expansive view*

In her description of the harms of subordinating speech, MacKinnon writes, “[a]ll the harms of [racist propaganda]—including stereotyping, objectification, deprivation of human dignity, targeting for violence, and terrorization of target groups—are harms of social inequality” (1995, 301). To this, she adds that social inequality:

is substantially created and enforced, that is, *done*, through words and images. Social hierarchy cannot and does not exist without being embodied in meaning and expressed in communications. [...] Segregation cannot happen without someone saying ‘get out’ or ‘you don’t belong here’ at some point. Elevation and denigration are all accomplished through meaningful symbols and communicative acts, in which saying it is doing it. (307)

What MacKinnon is here calling social inequality is what I am identifying as the third area of harm in my analysis of subordinating speech: its role in shaping the unjust group-based hierarchies that partially shape oppression. This area, I claim, has a much wider scope than what is plausibly captured by the attention given to the illocutionary force of subordinating speech acts by philosophers of language.

To be clear, as I see it, illocution is a useful analytical category for describing the force and immediate impact of many cases of subordinating speech. Chanting “Jews will not replace us,” broadcasting “All Tutsi are snakes and must be eliminated” on the radio, and posting “We don’t serve Muslims” in a business are all helpfully seen as words that *do something in the here and now*, and illocution aims to capture that. However, if we are to fully capture the role that words, speech, and language play in enabling and sustaining oppression, we must broaden our focus beyond illocution.

Harms arising from speech acts that are less immediate than threats or incitements can and do play crucial roles in setting the stage for the oppressive practices elaborated by Young and others. Not all of these are captured by illocutionary acts, and this is for good reason, as

illocution is meant to isolate the act performed in making certain utterances in certain contexts—that is, how “I do” becomes the act of marrying in certain context. Less immediate outcomes, even if they are foreseen, are standardly lumped into the perlocutionary, which comprises all the varied outcomes the speech act helped bring about. So, while marrying is illocutionary act, everything from upsetting your in-laws to making your parents cry are (some of) the perlocutionary effects of “I do”; so the example standardly goes.

Illocution is useful when the goal is to isolate the performative force attributable to *just* the speech act, absent any further contingent contributions. Yet, in the analysis of oppression and oppressive speech, our aim shouldn’t be limited to detailing the direct acts constituted by certain utterances, as much more deserves our attention. This is demonstrated by the expansive way in which the oppressive force of language is described by writers working outside of the narrow Austinean framework.

For example, in her chapter titled “Language,” bell hooks reflects on a poem by Adrienne Rich with the repeated line, “This is the oppressor’s language, but I need it to talk to you.” hooks goes on to say “I know that it is not the English language that hurts me, but what the oppressors do with it, how they shape it to become a territory that limits and defines, how they make it a weapon that can shame, humiliate, colonize” (hooks 1994, 168). She goes on to describe various ways in which language can have an unseen repressive force—such as the denigration of African American English Vernacular (AAVE), including its near total absence in academic circles.

Relatedly, there is Frantz Fanon’s remark that,

If a man who speaks pidgin to a man of color or an Arab does not see anything wrong or evil in such behavior, it is because he has never stopped to think. [...] To speak to black people like this is to express this thought: “You’d better keep to your place.” (2008, 20–21)

Philosophers of language concerned with the social and political dimensions of speech ought to resist frameworks that unduly restrict their focus, or risk losing sight of relevant phenomena. Because, when we ignore these broader aspects of the oppressive force of language, what we're left with is an impoverished understanding of reality. By marginalizing the ways in which language contributes to oppression, though only with further extra-linguistic support, the accounts of subordinating speech remain overly abstract. Moreover, seeing these cases as marginal misunderstands the importance language plays in these broader systems of oppression and relies on a division between speech and actions that ought to be more fully rejected, particularly by speech act theorists.

2.3.3. *Speech and action*

In my view, oppressive speech and oppressive acts are not separate practices best analyzed apart from each other but inform one another as together they constitute the realities of oppression, both in its abstract linguistic, as well as in its brute, material forms. In this respect I believe I am following MacKinnon, who argues:

Together with all the material supports for inequality, authoritatively *saying* someone is inferior is largely how structures of status and differential treatment are demarcated and actualized. Words and images, are how people are placed in hierarchies, in castes. They are the way social stratification is made to seem inevitable and right, how feelings of inferiority and superiority are engendered, how indifference to violence against those on the bottom of these hierarchies is rationalized and made to seem normal. (1995, 309)

As I understand MacKinnon's view, this is not to ascribe any mysterious powers to language for its ability to subordinate and harm. Rather, it describes a reciprocal relationship between the words and images used for and to oppressed groups and the non-linguistic oppressive practices they inform and rationalize. As she points out, this includes practices that silence the voices of

marginalized people, rendering their suffering invisible, while the ‘burdens’ of their oppressors are at the same time given a privileged place in most political discourse.⁶⁸ In this way, speech, words, and language both set the stage for violent oppressive acts, and also provide the lens through which these acts are interpreted. It is in this way civilian murder becomes ‘collateral damage.’

Seeing this reciprocal relationship follows naturally from understanding oppression in structural terms, where the meaning of individual acts is revealed through attention to the broader practice of which it is a part. For this reason, it is imperative for philosophers concerned with the role of speech in oppression to acknowledge the situatedness of that speech in a broader system of oppression that includes many different types of acts. As Lederer and Delgado put it: “the purpose of hate speech is the subordination of one people by another,” and while the “mechanisms of this subordination are different in every case,” in any event, “they usually include a complex, interlocking series of acts, some physical, some verbal, some symbolic” (1995, 5). Isolating the linguistic aspect would serve only to separate one feature of subordination from its larger web.

Violence, exploitation, material inequality, segregation and much more are the lived conditions of oppression. Philosophers concerned with how language enacts and maintains subordination ought to be concerned with these features, and their accounts of how speech oppresses ought to illuminate these conditions. However, as I argue below, the standard model of oppressive speech, which focuses on illocutionary aspect of utterances, is not up to the task of

⁶⁸ One striking example of this occurs in the way white male perpetrators of sexual assault are often viewed sympathetically and undeserving of any harsh punishment as it will leave too long-lasting a scar on an otherwise admirable life, while survivors are blamed for somehow bringing the assault on themselves. Kate Manne has recently given the name “himpathy” to this phenomenon (Manne 2017, Ch. 6).

illuminating the realities of subordinating speech. This is the case because these broad structural conditions most naturally fit among the perlocutionary effects. Yet, for the reasons elaborated above, Langton and McGowan are most invested in claiming oppression to be an illocutionary aspect of speech. In the next section I examine some problems this approach inevitably encounters and suggest that this general strategy is not as worthwhile as is assumed.

3. AMBIGUITIES IN THE ILLOCUTIONARY

Whether there exists a sharp line separating the illocutionary from the perlocutionary aspects of a speech act is a philosophically rich question. However, in what follows I want to bracket how this question might be addressed and answered in the case of non-subordinating speech acts. That is, I want to put aside what the right answer might be for morally-neutral utterances—if such a category even exists is a question I leave aside—and instead focus primarily on how this division plays out in specifically subordinating speech. While this approach has the potential drawback of allowing a lot of moral noise to interfere with our judgments about cases, I believe it makes up for this by keeping the relevant explanandum in view.

This approach has the additional benefit, I believe, of avoiding some of the pitfalls that existing models encounter, where instances of oppressive speech are analogized to more routine discourse contexts where a coherent conversation with clear contours and participants is the paradigm. This shouldn't be our starting point for analyzing subordinating speech, as it runs the risk of forcing oppressive practice to fit the philosophy of language, rather than the other way around.

3.1. *The Indeterminacy of the Illocutionary*

Addressing the question of the (in)determinacy of the illocutionary act requires returning to the distinction between causal and constitutive notions of harmful speech. Looking here reveals a tension in how theorists who claim this is the primary mechanism through which speech subordinates understand the notion of illocutionary harm. As I noted above, these theorists suggest that the main harm of subordinating speech is more *immediate* than some downstream effects it might cause—hence the greater focus on the illocutionary act rather than (subsequent) perlocutionary effect. In other words, we need not wait and see if some hateful utterance later results in an oppressive act like gender-based violence, the speech itself is an act of subordination.

At the same time, these theorists also hold that “speech can harm not just directly, such as by causing fear and anxiety in its targets, but also somewhat indirectly, by affecting the positions of groups to which those targets belong within the social hierarchy” (Maitra and McGowan 2012, 7). That is, speech sometimes harms not by being the direct cause of psychological pains, but by *indirectly* fixing facts “about the distribution of social power, including facts about who has this power, and who lacks it” (ibid.). The classic example, discussed above, is that of a legislator declaring that members of a minority group no longer have the right to vote. In doing so, they make it the case that this group now loses a power they once had, and now comes to occupy an inferior position in the social hierarchy. And this is apart and quite distinct from speech harming *directly* by itself causing psychological distress of some sort.

In the combination of these two claims, however, a tension emerges.⁶⁹ The illocutionary act is both the *immediate* act being done *in* saying those words; and it is also the act whose chief result is “somewhat indirect,” mediated by various institutions across time and space. To be clear, these two claims are not incompatible, but nor are they equivalent; and I will claim they pull in opposite directions.

To start, I believe this is the result of two distinct notions of speech *causing* harm functioning as the contrast to speech *constituting* harm in each case. First is the idea that speech can cause harm by infecting the hearts and minds of (some of) its audience and uses a long-term notion of cause. For example, anti-Muslim hate speech, in its propagandistic mode, spreads out from the original speakers and is taken up by some who then act on its message in violent ways. This account of speech causing harm is the ‘mentally-mediated’ variety, where the mind referred to here is that of perpetrators of violence, who are themselves not the speakers nor the victims.

The second account is the more direct causal variety, where speech can cause harm by *immediately assaulting* its target, most obviously by causing psychological distress like fear, anxiety, and much more. And while the full effects of the assault might take days, weeks, or a lifetime to fully manifest, the idea here is that the utterance functions as the moment of assault.⁷⁰ This is one idea of how speech causes harm, and it is the type of harm I surveyed in section 2.1.

Both notions of speech causing harm are contrasted with the claim that speech itself constitutes a harmful—indeed oppressive—act. Compared to the first notion, the constitution view says speech can also harm *immediately*. And compared to the second notion, the

⁶⁹ It is worth noting how these different understandings may be due to the former view—the immediate harm—coming most explicitly from Langton, while this latter approach—the indirect harm—coming from Maitra and McGowan. To be sure, the accounts offered by these three theorists are not interchangeable. Nonetheless I see a similar underlying commitment to the primacy of illocution in each, and so treat them similarly.

⁷⁰ See Tirrell (2017) for a conception of harmful speech as functioning like a toxin. And like toxins, there is variability in terms of how long a toxin may remain latent before its full effects are known.

constitution view says that speech can also harm *indirectly*, by fixing normative facts rather than like an assault on the physical person.

I bring up these two contrasts to point out how, on the standard illocutionary framework, both causal notions are captured by the perlocutionary effects of a given speech act. This is the case because both are aspects of how speech can evince different “states of mind and behavior in ourselves and others” (Bauer 2015, 58). Both the immediate pangs of fear in the target of hate speech and the gradual solidification of hateful attitudes in the bigot fall in this category, and both are distinct from the normative change described by the illocutionary act.

This wide breadth of the perlocutionary is no accident. It is, so to speak, a feature, not a bug. In Austin’s example of the one man saying “shoot her” to the other, the realm of the perlocutionary covers everything from the second man’s *surprise* at the first man’s utterance, his being *persuaded* to shoot, and the woman being shot. To this we might add her family’s grief, the trial for their crime, both men’s guilt, their imprisonment, and so on. While these effects are a motley crew, they are all distinct from the illocutionary act, namely, the first man’s *advising* that the second man shoot her.⁷¹ Seen in this way, it is obvious that the perlocutionary effects encompass a lot, while the illocutionary is intended to narrow our focus to the particular act inherent in the utterance itself.⁷²

This is a natural consequence of the Austinean picture, where the category of the perlocutionary is meant to “introduce the idea of extra-linguistic or incidental consequences of speaking—of, as it were, *further* things that are done for which the conventions attaching to

⁷¹ Or was he *ordering* her? Even though illocution is relatively narrow, it is still open to difficult ambiguities.

⁷² Though I say particular act, it need not be a singular act, and the idea that a single utterance can encompass a number of acts is not incompatible with the illocutionary being a narrow category. That is, speech act pluralism doesn’t threaten the main point I hope to be drawing out here.

forms of speech do not by themselves provide” (Hornsby and Langton 1998, 24). And it is this understanding of the (narrow) illocutionary act and the (expansive) perlocutionary effects that leads many speech act theorists—including Austin—to focus almost exclusively on illocution, which is considered more analyzable because of its relative constraint. But as I will begin to argue now, this leads to a faulty approach to subordinating speech.

Let me begin with another example: I *invite* you to dinner. You are caught off-guard and *refuse* my invitation. I am embarrassed. Here, my *invitation* and your *refusal* are properly understood as illocutionary acts, because they are acts determined by linguistics, semantics, and convention. You’re being surprised, though, and my embarrassment are perlocutionary effects, along with many other things, as these are incidental consequences not fixed by language, but rather by other psychological, social, and extra-linguistic facts. When all goes well, this is how the analysis works.

But things aren’t always so neat. Sometimes, the status of an illocutionary act is in doubt. There exists a substantial literature on the complicated status of several important speech acts when the audience’s uptake is in some sense not optimal. Can we say that a speaker actually promised, warned, or refused if their audience doesn’t take them to be promising, warning, or refusing?⁷³ According to Austin, a warning is at least in some sense “unhappy” if the hearer to

⁷³ For philosophers like Langton and Hornsby who follow Austin and take audience uptake to be a necessary component of the illocutionary act, this is a crucial topic. It opens up a key line of analysis and argumentation: namely, if audience uptake—that is, according to these writers, the hearer’s recognition of the speaker’s intent—is required for successfully pulling off an illocutionary act, then the man who doesn’t understand the woman’s ‘no’ as a refusal makes it the case that the woman can’t actually refuse, and she is, in effect, (illocutionarily) silenced (Langton 1993; Hornsby 1994; Hornsby and Langton 1998). And, according to Hornsby and Langton’s argument, women can be illocutionarily silenced by a patriarchal society in general, and pornography in particular. However, between the woman’s ‘no’ and the man’s (psychological) response, it remains unclear whether any illocutionary act has occurred at all. It is ambiguous.

whom it is directed doesn't understand the speech act to be a warning—perhaps instead taking it to be a joke.

While this is an issue for many types of speech acts, the problems multiply in the case of subordinating speech. A straight woman mutters “gross” when her two gay friends kiss. They react with surprise and alarm concerning her homophobic comment. She counsels them not to be so sensitive, as she was “just joking.” But was she? Is her comment an ‘unhappy’ joke, or an oppressive judgment, or perhaps both? When it's not obvious what illocution she performed *in saying* her speech act, it's sensible to look to the further effects her utterance may have caused to aid our analysis. But as sensible as this is, it clearly shows the dependence of illocution on perlocutionary effects.

This is most evident when considering the notion of ‘accommodation,’ which plays an important role in many recent accounts of subordinating speech.⁷⁴ Here, the full force and even type of a given speech act is partly determined by whether an audience lets a speaker's words go unchallenged, or whether they engage in counter-speech. In between the audience's silent accommodation or blocking speech act, however, the original utterance exists in an ambiguous state. Saray Ayala and Nadya Vasilyeva make this ambiguity explicit when they write: “We picture a sort of limbo in which the conversational score is updated after speaker's utterance and the permissibility facts are to be enacted unless something is done to stop it.” From this they conclude that the “interlocutors' move is constitutive of speaker's speech act” (Ayala and Vasilyeva 2014, 6–7).⁷⁵

⁷⁴ Some recent papers where something like accommodation plays a key role in the overall account include Maitra (2012), Langton (2012), Langton (2017; 2018a; 2018b), Ayala and Vasilyeva (2014) and more. I return to this topic throughout this dissertation, mainly in Chapters Two, Three, and Four.

⁷⁵ Discussions about the different effects of ‘calling out’ versus ‘calling in’ do much the same work in revealing the importance of audience reaction to the status of potentially subordinating speech.

As plausible as this may be as a description of how subordinating speech functions in situ, it reveals a tension in the underlying account. In short, if the type of speech act that a given utterance constitutes is only established *after* being interpreted, and perhaps blocked or accommodated by its audience, then the illocutionary act depends in part on the perlocutionary effects, and the divide cannot be so strict. This poses a serious problem for accounts that focus on the illocutionary act as rigidly distinguished from its perlocutionary effects, as Rebecca Kukla (2014) points out. She argues:

on [Hornsby and Langton's] account, performative force is not effected in the act of speaking, but rather partially constituted by a wholly separate, contingent subsequent event, namely the audience's recognition of the speaker's intention. But this recognition is a perlocutionary effect of speaking, and hence the performative force they are talking about does not seem to be illocutionary after all. (2014, 454)

How serious this problem is depends on one's understanding of the illocutionary.⁷⁶ For some, this is simply a straightforward consequence of Austin's theory, and shouldn't be taken as any sort of objection (Hornsby and Langton 1998). For others, like Daniel Jacobson, this demonstrates how an implausible understanding of Austin holds "the performance of an illocutionary act hostage to the perversity of one's audience" (1995, 74). And while Jacobson rejects this view, he nonetheless admits that on his own approach to speech act theory, it is "no simple matter to determine when a given illocutionary act has been brought off" (ibid.). Because a speaker always requires *something* of their audience to successfully perform an illocutionary act, there appears to be an indeterminacy inherent to this category.

I believe this represents a significant problem for accounts that prioritize the analysis of constitutive harms embodied in illocutionary acts. This is the case, because even though authors

⁷⁶ Maitra argues that the inability to properly demarcate the illocutionary from the perlocutionary is a problem that ultimately undermines Hornsby and Langton's conception of silencing. See Maitra (2009).

like McGowan explicitly reject attending to what speech *causes*, illocutionary acts in fact require perlocutionary effects to be the acts they are. As such, causal effects must come into consideration, and the distinction is not as sharp as is suggested. In the next section I demonstrate how this is not simply a problem for the coherence of the concept of illocution—which is not my primary concern—but also for the general analysis of subordinating speech. That is, rather than clarifying the core features of subordinating speech acts, the focus on illocution in fact pushes the analysis away from crucial features of the overall speech situation.

3.2. *Felicity Conditions and ‘Paradigm Cases’ of Subordination*

Generally, if we want to know whether a given illocution occurred, we need to look to what Austin calls the “felicity conditions” of that illocution. These are the features that determine whether a given speech act was successful or not—or “happy” or “unhappy,” to use Austin’s terminology (1962, 14). For example, the reason why some “low type” can’t successfully christen the ship ‘the *Generalissimo Stalin*’ even though she swoops in, smashes the bottle and utters the right words, is because not all of the necessary felicity conditions are in place. Namely, only speakers with the right entitlements—the right kind of authority—can christen the ship. It’s not just anyone with a bottle. I’ll return to the issue of speaker authority in subordinating speech in more detail throughout the dissertation. Here, though, I want to bring out and question the basic idea of identifying the felicity conditions of subordinating speech acts.

The difficulty of identifying the felicity conditions of a given speech act type can vary from relatively straightforward to seemingly impossible. On the easier end there are the highly structured speech acts typical of formal organizations. What specific conditions need to be in place to call a meeting to order, or swear someone into office, is often explicitly written down in

a rule book somewhere. But in less formal settings, the felicity conditions are harder to identify, and disagreement about whether a speech act was successfully pulled off, or even which particular act was performed, is common.

I tell you “I’ll be there!” when you invite me to the bar you’re be going to tonight. Did I make a promise, state my preferences, or just express an expectation of myself? Maybe you and I disagree, especially when you show up, but I don’t. “In situations of disagreement,” Langton tells us, “the disputed illocution usually falls short of the paradigm case for the given locution.” In the paradigm case, she continues, “one knows just what the felicity conditions for the given illocution are, and one knows that they are all satisfied” (1993, 308). For many illocutionary acts, the paradigm cases wear their performativity on their sleeves quite explicitly. Had I said, “I promise to be at the bar later tonight,” we would likely both agree that I made a promise, even if I didn’t fulfil it in the end.

For the class of utterances Austin names “explicit performatives,” the speaker makes their action plain through grammar and syntax. The main examples are often ceremonial, as in: “I hereby christen this ship *Curly’s Gold*”; or “By the powers vested in me, I hereby I pronounce you married”; or “As the scheduled business has all been attended to, I hereby adjourn this meeting.” These illocutions, always in the “present first-person singular indicative,” and often including the telltale word ‘hereby,’ offer one image of the paradigm that gives a particular illocution its felicity conditions.

But things need not be so formal. If the chair of the meeting says, “meeting’s over,” or “we’re done here,” it would likely the same effect as her more explicit utterance. And note that any version of her utterance will have a different effect than that of an onlooker saying very the same words—making the same locution. So, surface grammar and locutionary content only offer

rough clues to the felicity conditions of a given illocution. This is a straightforward result of the observation that the same locution can do a great many different illocutions, and vice versa. The sentence “It’s a bit too salty, isn’t it?” could function as a genuine interrogative, or, depending on the circumstances—the “total speech situation”—come off as an assertion, a request, an order, or even a devastating insult.

In the case of subordinating speech, things get even more complicated. Langton, for her part, recognizes this difficulty and takes on the challenge it presents in her application of speech act theory to pornography. In arguing that pornography constitutes the subordination of women, Langton acknowledges that this is a disputed interpretation of its illocutionary force, and suggests that it is analogous to situations where people might disagree whether an utterance really constituted a *marriage* (perhaps because the officiant lacks some important credentials) or if the utterance “shoot her” really constituted *ordering* (rather than *advising*, as the speaker maybe intended). Nevertheless, she argues that, much like in these situations, we can make progress on deciding what illocution occurred by asking whether the disputed utterances match closely enough the paradigm cases where the felicity conditions are both understood and successfully met.

When illocutionary force is in dispute, Langton claims we can make the case for a particular illocution having occurred in at least three ways. First, we might claim that “vagueness or ignorance notwithstanding, *some* felicity conditions—important ones—are satisfied, and that is good enough.” Second, we may examine whether the “*uptake* appropriate for the claimed illocution has been secured.” Or, third, if “a speech act’s effects are best *explained* by supposing that it has a certain illocutionary force” we may then conclude the supposed illocution in fact occurred (1993, 309). Langton acknowledges that “all three ways of arguing are fallible, and

they come in ascending order of fallibility,” but suggests that “each of the three, or some combination of them, may be useful, depending on the evidence we have” (ibid.).

To make progress on this question, then, Langton uses the example of the apartheid-era South African legislator as the paradigm case for subordinating speech.⁷⁷ This example is worth exploring in more detail. First, let me put aside one major dis-analogy that Langton claims can be overcome. This is the fact that the South African legislator, but not the pornographer, occupies a formal position of authority that gives their speech more force than it otherwise might have. Langton claims this question is “at the heart of the controversy” and I put this aside only to take it up more fully in my next chapter, which directly addresses this Authority Problem for subordinating speech.

Even putting this problem aside, however, other issues remain. My general concern here is with the philosophical tendency to abstract away from the broader contextual features of an utterance with the hope of isolating (and therefore investigating) some supposedly more central conditions of a subordinating speech act. In Langton’s case, this is demonstrated by which features of the apartheid-legislator example she draws out and puts to use.

Regarding her second and third strategies noted above, Langton suggests that the speech acts of apartheid “are speech acts that achieve a certain uptake: they are taken to be verdictive and exercitive acts (though not all hearers will take them to be subordinating acts).” Furthermore,

they are illocutions that have a pattern of perlocutionary effects on the beliefs and behaviors of the population: whites believe blacks to be inferior, believe discrimination against them to be legitimate, and believe them to have fewer rights; whites discriminate against blacks, and blacks stay away from polling booths. (1993, 310)

⁷⁷ To be precise, Langton does not explicitly claim the apartheid speech act is the paradigm. Specifically, she writes: “I have not tried to say exactly what the paradigm for subordination is, but I have suggested that the speech acts of apartheid offer a clear example” (1993, 310). So perhaps it is better to say she takes these to represent *a* paradigm, rather than *the* paradigm.

That is, according to the (admittedly fallible) method of argument to the best explanation, “part of the explanation for whites’ discriminatory behavior is that such behavior has been legitimated by law. Part of the explanation for blacks keeping away from certain areas is that they have been ordered away” (1993, 309).

It’s important to note the modesty of Langton’s claims here. First, it is only *some* hearers who take the speech acts of apartheid to be acts of ranking and legitimating (verdictive and exercitive acts, respectively). Second, these utterances serve as *part of* the explanation for the effects in question. They aren’t the whole explanation, but one, she believes, is worth isolating our attention on.

This argument is thus an instance of the idea, already explored above, that a given “pattern of perlocutionary effects”—in this case, the “beliefs and behaviors of the population”—are there because of what the speech is, as an illocutionary act. However, here we run into the problem discussed in the previous section, raised by Kukla. Namely, if it is only because of the perlocutionary effects of speaking that a speech act has a certain force, then it doesn’t seem to be, strictly speaking, an illocutionary act (Kukla 2012). Or at least, the divide between perlocutionary and illocutionary—and the greater attention given to the illocutionary—ought to be suspect.

It was because of this problem that we were first led to examine the felicity conditions of a type of speech act. As for the important felicity conditions of subordinating speech acts, Langton explicitly discusses only speaker authority. Yet, some of the important—indeed necessary—contextual features aside from the legislator occupying a position of authority likely include other facts, even in this paradigmatic example. For instance, the existence of a military/police force to enforce the government’s legislation is required. There is also the fact

that a powerful anti-Black minority agrees with the laws in question, and crucially, most likely did so on the basis of beliefs they held *before* the legislator's speech act. There is the existence of negative stereotypes about the Black majority. There is the selective media coverage that exacerbates these stereotypes. These features all appear just as crucial to the legislator's felicitous illocution as their formal role in the overall oppressive apparatus.

We can see this simply by imagining what would follow had the legislator instead uttered "whites are not permitted to vote," from the same position of authority. My intuition is that the result would be widely different. This suggests that the felicity conditions necessary for subordinating speech are more expansive than Langton's initial gloss brings out, extending beyond mere speaker authority and into the realm of social and political context within which the utterance occurs.⁷⁸ On Langton's analysis, though, the connection to oppressive structures remains implicit, but it does more work than she acknowledges. Indeed, it is only because of these implicit connections to the broader political context that her example is as compelling as it is. Once we remove them from the example, it loses its force.

In short, the strategy of narrowing our focus to some 'core' features of subordinating speech acts is fraught with difficulties. It requires identifying the felicity conditions of subordinating speech. But this approach is problematic because it is doubtful we could, even in principle, identify paradigmatic instances of subordination. Because of its totalizing and structural nature, subordination shapes both the macroscopic and microscopic features of one's life under its thumb. Attempting to isolate paradigmatic cases of subordination—and

⁷⁸ Ramsey (2013, 112–13) makes a similar observation about Langton's legislator example, using the example of "people under 21 are not permitted to vote." And, in a similar vein, Tirrell (2018b) notes the larger context that renders a judge's sentencing utterance effective (13); and makes explicit how "[i]nstitutional settings often mask the audience's role in successful speech acts. [...] Even speech that feels like the exercise of pure authority nevertheless relies crucially on hearer uptake" (21).

subordinating speech—inevitably pushes important and complicated dynamics aside in favor of what’s more observable. In Langton’s case, this means isolating speaker authority as the sole felicity condition of subordinating speech. But, by lingering on the example, we’re led to the thought that there is so much more at play. And if this is a problem for the paradigmatic case, it is only more so in the more mundane and everyday cases that litter the real life of subordinating speech.

In sum, Langton, along with other philosophers such as Maitra and McGowan, pursues a narrowing approach by using illocution as the central concept.⁷⁹ Yet, if we take seriously the thought that our analysis must address “the total speech act in the total speech situation” (Austin 1962), we are led to believe that this task isn’t only difficult, it’s impossible, and moreover obfuscating, as it pushes aside key features that an adequate analysis must include.

3.2.1. Ideal and nonideal (speech act) theory: A short detour

My criticisms from the previous section can be strengthened when we consider further idealizations implicit in some strains of speech act theory. Parallels that exist between the ideal/nonideal theory debate in political philosophy and the concerns I am raising about an illocutionary-focused approach to subordinating speech bring this out well. At first blush this might seem a bit odd, since the topics that Langton and others use speech act theory to illuminate are undoubtably nonideal. This is true. Nonetheless, the methodology of speech act theory, as employed by Langton and others, engages in some unhelpful idealizations that undermine its applicability, and ultimately hinder its ability to analyze the problems at issue.

⁷⁹ It is worth noting, though, how each at the same time notes the difficulty of such a pursuit, with Maitra being, in my eyes, the most skeptical of its ultimate plausibility.

As Mitch Green observes, “Speech act theory traditionally idealizes away from many aspects of a communicative situation, including those in which one speaker is marginalized or another is given undue authority” (2016). Mary Louise Pratt makes a stronger claim when she describes the “ideological dimensions of speech-act theory,” which tend to designate “much of what people do linguistically as lying outside the norms of their language” (1986, 70). And while both are referring mainly to more traditional accounts like those of Grice, Austin,⁸⁰ and Searle, we can note how this tendency to idealize and abstract has not been fully overcome, even in cases of subordinating speech.

In her “The Ideology of Speech-Act Theory,” Pratt notes how some assumptions in speech act theory can be taken too quickly for granted. These include the idea that one-to-one speech functions as the paradigm of all other linguistic acts, and the assumption of an underlying unified (true) subject as the seat of beliefs, desires, intentions, etc., against which the speech act is to be assessed. These assumptions make cases where there are many speakers and hearers, and where a speaker speaks for, or through, others seem abnormal. Yet, this is odd, because it is no way abnormal for a person to be speak “as a member of some collective, or as a rank in a hierarchy” (Pratt 1986, 63). Most people do it every day as part of their job.

Because of a set of assumptions and absences, Pratt claims that speech act theory tends to normalize a particular subset of actual communication, and these “‘normalized’ forms of expression invoke a harmonious and homogeneous social world that is quite different from any known social formation” (1986, 69). The problem then, is not only one of omission, but

⁸⁰ Though it is worth noting that Pratt distances her critique from Austin and is more concerned with those who came after him, namely Searle (1976; 1994).

ideological idealization.⁸¹ Pratt's critique is thus reminiscent of Charles Mills argument in his "Ideal Theory as Ideology"—despite her article appearing twenty years earlier. There Mills claims that what "distinguishes ideal theory is the reliance on idealizations to the exclusion, or at least marginalization, of the actual" (2005, 168).

In the case of Langton, this problem emerges in terms of the paradigms against which she analyzes subordinating speech acts. Aside from the specific issue the nonexistence of a paradigm of subordination poses for identifying felicity conditions—considered above—there is the further problem that Langton's apartheid example treats the legislator's speech as that of a single speaker. She speaks of *the* legislator, after all. In this way, the legislator's role as a mouthpiece for the state is rendered negligible. But of course, it matters greatly that their utterance is not attached strictly to their own intentions, beliefs, etc., but that of an office with specific powers.

This problem extends when considering more 'everyday' occurrences of subordinating speech as well. Because of the ideological assumptions inherent in the paradigm, we might be led to investigate whether the particular speaker holds specific intentions or beliefs—for example, the intention to discriminate, or beliefs about the inferiority of people of color, etc.—and whether they have authority in a way similar to that of a legislator. This approach, however, takes the problem of oppression and examines it through an overly individualized lens. It matters very little what the precise intentions or beliefs of a given speaker are when they say things that are reasonably interpreted as hateful, often by using the same words and phrases as out-and-out

⁸¹ As she elaborates, "these lines of normalization are ideologically laden. Among the cases systematically marked deviant, one invariably finds all ludic kinds of expression (language as play rather than productive work); some forms of expression primarily associated with women, such as gossip, small talk, or euphemism; and other forms like circumlocution, indirectness, or deliberate ambiguity, that are associated with communication across hierarchy and across lines of conflict. [...] There are good reasons to be dissatisfied with a theory that designates much of what people do linguistically as lying outside the norms of their language" (1986, 69–70).

bigots.⁸² And by using a legislator's speech as the yardstick against which other potentially subordinating speech acts are measured, we're led to a conception of speaker authority where it is easily recognized and identified, though we have very little reason to suppose this will be the case generally.

In general, one problem posed by positioning a particular speech act as the paradigm of subordinating speech—in this case, the apartheid legislator's speech—is that it implicitly endorses a hierarchy of subordinating speech. Oppressive speech acts are measured by how closely they match the paradigm. This is like the ideal theorist's approach to justice. Once the ideal is identified, injustice is found in deviations from that ideal—and the further the deviation, the greater the injustice. And much like the nonideal theorist's critique of that general approach to political philosophy, we can similarly recognize that this approach is not the only game in town.

Rather, we can adopt a nonideal speech act theory where we begin from the actual, and describe the features found there in detail. Then, from an adequate description of the lived experience of subordinating speech we will have the resources available to develop appropriate models and theories. Beginning from a conception of the paradigm of subordinating speech serves no purpose here, when there is little reason to suppose there is but one type of subordinating speech act. Rather, it is a diverse and multiply realizable category, and we ought not to force any description upon it divorced from actual experience.

⁸² I elaborate on this in Chapters Three and Four.

3.3. Assessing the Illocutionary Account

With the preceding description of some problematic aspects of a focus on the illocutionary act, we can return to the question of whether the attention philosophers give to it in accounts of subordinating speech are worth these drawbacks. That is, do the analytical insights offered by the illocutionary account of oppressive speech outweigh the potential hurdles the account must overcome. Above I noted two advantages the claim that speech constitutes harm—as an illocutionary act—was meant to secure. These are (1) attributing responsibility directly to the speaker of harmful utterances, and (2) drawing attention to normative changes in social status rather than only violent acts or psychological distress. I will take each of these in turn.

3.3.1. Attributing responsibility to the speaker

One benefit of identifying the illocutionary act as a harmful act in itself is that it arguably follows directly from this that the speaker is in some sense responsible for the harm. Rather than only being a source of motivation for hate-based violence and discrimination perpetrated by others, the speakers of subordinating speech acts themselves are guilty of performing harmful (illocutionary) acts. This side-steps the ‘mentally-mediated’ defense of absolving the speaker from harm by noting that harm is not found only in the violent physical outburst of bigots, but also in their verbal acts. By fixing attention on the illocutionary harm, speakers themselves are put on the moral hook for their words. This move does well to bolster the claim that people are—or ought to be—responsible for their words, and that common excuses like ‘mere words’ or ‘just trolling’ are inadequate.

However, with the previous section's analysis in mind, I want to argue that things are not so simple. The problem, in brief, is that if the goal concerns assigning responsibility, then it matters what model of responsibility is being presupposed and invoked, and under the illocutionary model, that is primarily an individualistic, blame model of responsibility. Yet, as Iris Marion Young (2004; 2006; 2011; 2013) and others have argued, the systemic nature of oppression calls for a distinct model of responsibility beyond the traditional blame model.

It's useful to note that the blame model of responsibility functions mainly by isolating the blameworthy individual(s), and therefore has the tendency of absolving others. This is exactly why the 'mentally-mediated' defense is invoked by those who utter hateful speech. They might admit (publicly, at least) that an act of racist violence was a grave moral wrong but claim that the blame lies solely with the perpetrators of the act. When this plea is successful, it serves to absolve the speaker of any responsibility. It works (when it does) by characterizing the claim that a single utterance has the power to subordinate on its own as nonsense. And they do so by comparing this to the more tangible harm of physical violence.

So, while the illocutionary model may do well to get the speaker on the moral hook, it does so at the cost of isolating their speech from its wider context, which is picked up and ridiculed by critics. Thus, there is a severe cost to this approach, and one that should give us pause. As Matsuda notes, "[p]art of the special harm racist speech is that it works in concert with other racist tools to keep victim groups in an inferior position" (1993, 39). To properly capture this fact, we cannot separate the harmfulness of subordinating speech from the harms of oppression more broadly, but the focus on illocution runs this risk. Because of this, isolating speech from further facts of oppression is a tactic that ought to be resisted, but it is implicit in the model that foregrounds the illocutionary act apart from its complete context.

Before moving on, it's useful to consider briefly the motivation one might have for tying moral responsibility primarily to illocutionary acts, and only perlocutionary effects secondarily, if at all. The worry is that perlocutionary effects are, in an important way, out of the hands of the speaker, and so beyond their control. This is certainly true when we consider how wide the scope of the perlocutionary is meant to cover, at least according to the standard Austinean conceptualization. And if the domain of the perlocutionary is beyond the control of the speaker, it seems to follow this is not apt area to locate responsibility—hence the move away from causal accounts.

Yet, there is some disagreement about the proper scope of the perlocutionary, and therefore its potential analytical helpfulness. Some recent work that attempts to analyze speech acts primarily in perlocutionary terms seem to be, in my opinion, implicitly pushing back against the interpretation that the perlocutionary effects are simply anything and everything that causally follows from an utterance, and thus totally unhelpful.⁸³ I see myself as fitting within this camp of speech act theorist who resist the idea that this approach to language is solely or even primarily concerned with illocution.

For example, in aiming to expand this aspect of speech act theory, Stanley Cavell remarks that while it's true “the perlocutionary effects of an utterance may be as various as the motives for speech, and, as with any human action, an utterance will have (in general) intended and unintended effects and consequences.” But he goes on to add, “to know what perlocutionary acts I am *liable* for ‘bringing off’ is part of knowing what I am doing and saying, or am capable of knowing and saying (2005, 174). That is, it is not entirely a mystery what effects our speech acts are likely to have. And in many cases, our very reason for saying something is to bring about

⁸³ See, for example, Saul (2018a) and Hominh (2016), for two accounts that make this goal explicit.

certain effects. To ignore this aspect of speech is to push aside as irrelevant many of the key characteristics of subordinating speech, namely, its effects on the wider community.⁸⁴

As the discussion of section 2 is meant to show, the effects of hate speech and other forms of oppressive utterances are wide ranging, and to reduce these to a narrow set of features as the illocution model does risks downplaying the experience of the targets of such speech. Furthermore, in many cases, hate speech is often intended to have these effects. That is, the speakers of hate propaganda aim not only to *advocate* for racial hierarchy in an illocutionary sense, but also to *convince* their audience of the need for action. This is a clear perlocutionary goal that needs to be accounted for in our analysis. It is in service of these goals, and much more, that they often seek as many platforms and as much coverage as possible—seemingly working according to the belief that there’s no such thing as bad publicity⁸⁵—and have their eye on their wider audience.⁸⁶ Moreover, fear, anxiety, and a sense of isolation are not well captured in illocutionary terms. But the account of responsibility implied by the focus on illocution narrows the blameworthy acts too much, and fails to capture this wide range of harms—both causal and constitutive—that ought to be kept in view.

⁸⁴ Despite the general focus of her account, Langton herself makes this point when she writes: “Besides expressing feelings, hate speech *provokes* feelings. It invites an emotional response, as well as a cognitive and practical one. It tells someone what to feel, as well as what to believe, and what to do. Its effects—its perlocutionary goals, in Austin’s terms—include hatred, for some hearers, pain and fear for others” (2017).

⁸⁵ Indeed, one section title of the leaked style-guide for the neo-Nazi website *The Daily Stormer* was “All Publicity is Good Publicity.” See Phillips (2018) *The Oxygen of Amplification* for an analysis of how far-right groups exploit journalists in order to increase the reach of their messaging. For a liberal defense of the practice of ‘no platforming’ consistent with strong freedom of speech, see Simpson and Srinivasan (2018).

⁸⁶ Stanley Cavell notes the importance of the second person in perlocutionary acts, which he contrasts with illocutionary acts in the following: “As Austin insists, in the case of illocutionary acts, even though it may be only implicit in various grammatical forms, ‘The ‘I’ who is doing the action does [...] come essentially into the picture.’ So I might comparably say: In perlocutionary acts, the ‘you’ comes essentially into the picture” (Cavell 2015, 180).

3.3.2. *Normative changes in social status: The construction and maintenance of hierarchies*

To be sure, normative changes occur via speech acts. But, in the case of subordinating speech in particular, it is worth asking what sustains those changes. Between friends, one making a promise is often all it takes to normatively bind one to the other, and we recognize this when we say the promisee can reasonably hold the promiser to their word. It is in these cases that, as Austin is fond of saying, our word really is our bond.

Yet for oppressive speech—and oppression more broadly—the normative story is much more complicated. This came out in the above discussion of the omissions in taking the apartheid-era legislator’s utterance as a paradigm of subordinating speech. While I agree the legislator’s speech functions to rank Black South Africans as inferior and unjustly deprives them of their rights, it is not speech alone that kept Black South Africans oppressed. The enforcement of these laws, and the extra-legal violence that occurs alongside the law—often with its implicit approval—constitute other barriers to equality for subordinated groups.

While linguistic acts are one route by which racial (and other) hierarchies are established, justified, and spread, our understanding of these acts is enhanced when we put them in their proper place, alongside the other, nonlinguistic acts that back up these words with force. This, of course, includes violence, as being subject to violence simply because of one’s membership in a group is one hallmark of oppression (Young 2011).⁸⁷ Putting subordinating speech acts alongside non-verbal forms of harassment, intimidation, and violence therefore offers a clearer picture of how the overall pattern of these practices function to create and sustain oppressive hierarchies.

⁸⁷ Consider also Matsuda’s remarks that: “Violence is a necessary and inevitable part of the structure of racism. It is the final solution, as fascists know, barely held at bay while the tactical weapons of segregation, disparagement, and hate propaganda do their work” (Matsuda 1993, 24).

While I doubt Langton and the other philosophers I discuss would deny any of this, I maintain that the illocution-focused model they defend has the effect of isolating speech from its background. By analogizing oppressive speech to other illocutionary acts, they neglect how subordinating speech is unlike those other acts. John Michael Ramsey expresses this point nicely when he writes:

The force of oppressive speech is not explained by or exhausted by the illocutionary force of a speech act, as is the force of a promise-utterance or warning-utterance. Instead, the force of oppressive speech originates in and is explained by the background, normative practices of which the utterance is a move. (2013, 193)

In other words, we must look to the background, along with the other practices of subordination, if we are to understand the force of subordinating speech. This would be an improvement upon a view that locates its subordinating potential in terms of a single act—as the illocutionary model is apt to do.

My view recognizes subordinating speech as (one) part of a broader *practice* of subordination. It is for this reason I took the time to discuss the wide range of harms attributable to speech and situated this within an account of oppression. Doing so highlights subordinating speech acts as part of broader practices of oppression that includes structural violence, voter disenfranchisement, epistemic injustice, the wealth gap, and more. This also offers a better picture of subordinating speech even when we focus solely on speech. In an insightful footnote, Maitra recognizes this when she sketches an alternative view of understanding the claim that speech constitutes subordination. According to this alternative:

to say that speech can constitute subordination is to say that, under certain circumstances, the practice of producing speech of the kind in question is a subordinating *practice*. On [this interpretation] the practice of producing racist hate speech (for example) may be subordinating, even though each particular act in that practice isn't a subordinating act. (2012, 98n)

Maitra only discusses this alternative view of the constitution thesis once more, where she briefly suggests that it could assist us in our understanding of authority and responsibility for subordinating speech (Maitra 2013). There, she also notes the further advantage “that such a re-framing might avoid controversies over how to understand illocutionary acts such that subordinating falls into that category” (2013, 92n). This is the controversy that I have explored throughout this chapter. And as I have been suggesting, we do well to move away from the illocutionary act approach as our primary understanding of subordinating speech. This approach isolates speech acts from their broader context too much, and, in the case of subordination, this is unacceptable cost.

4. CONCLUSION

In this chapter I have offered a description of the expansive harms of subordinating speech and argued that the philosophical focus on illocution as the main mechanism by which speech subordinates is unfounded. This chapter had two main goals. One goal was to provide a comprehensive account of the plausible harms of subordinating speech. Any attempt to do so will likely fail to fully capture every aspect of such speech, and my account is abridged in many ways, I’m sure. Yet I hope that by marking out three distinct areas of such harms—as they are approached by writers with different theoretical commitments—I have both conveyed the wide scope of these harms and shown how they are inextricably related.

The second main goal of this chapter was to push back against a strain in social philosophy of language that privileges illocutionary acts as the proper lens for analyzing subordinating speech. I have argued that this standard neo-Austinean framework fails in a few key ways, namely, by isolating subordinating speech acts from the broader practices within

which they take part. In the next chapter, I follow up on an important topic that was side-stepped here: the question of authority in subordinating speech.

CHAPTER TWO: SPEAKING WITH (SUBORDINATING) AUTHORITY⁸⁸

1. INTRODUCTION

For those who are members of disadvantaged social groups, being the target of hate speech can be a regular experience that can lead to many physical and psychological harms—as I discussed in Chapter One. Yet, as we saw, many agree that the destructive element of hate speech extends beyond its ugly downstream consequences. That is, beyond being the cause of so many harms in the lives of its targets, hate speech itself seems to be a particularly harmful act. Indeed, along with many others, I see it as a *subordinating* act, one that marks its targets as inferior in an unjust act of ranking. Getting clearer about how this ranking occurs—indeed, how it is even possible—however, is a difficult task.

One problem arises from noticing how, for many speech acts, the speaker needs a special kind of standing to be able to perform them at all. That is, some speech acts require a particular entitlement, or authority, to be successfully carried out. This is a familiar idea. In everyday life we do many things with our words: we make promises, apologies, bets, and threats, and so on. Among these acts, some require a certain standing to accomplish successfully. To call a strike in baseball, you must be the umpire; to declare war in the USA, you must be the president. Yet, the type of standing required in each case is quite distinct. The president's saying 'strike' lacks the same force that my own utterance does. Neither one of us can make a ball a strike with our utterances, because neither of us has the *right* standing.

⁸⁸ This chapter is adapted from my paper "Speaking with (Subordinating) Authority," which was published in *Social Theory & Practice* 42 (2): 240–257 (2016).

For subordinating speech, this idea of appropriate standing leads to a challenge that has been called the Authority Problem. According to the Authority Problem, for speech to *constitute*—and not just cause—subordination, it must be spoken by speakers who occupy recognizable positions of authority. Yet, in the case of ordinary instances of hate speech—that is, hate speech uttered by the ordinary person on the street—speakers would seem to lack the required authority by definition, and so their words necessarily cannot subordinate.⁸⁹ On this view then, *ordinary* hate speakers—as we might call them—therefore cannot subordinate, simply because they are just that: ordinary.

In “Subordinating Speech,” Ishani Maitra (2012) defends the claim that ordinary instances of hate speech can sometimes constitute subordination. Her strategy in responding to the challenge posed by the Authority Problem is to admit that authority is indeed required for subordinating speech, while also arguing that “there are more ways in which a speaker can come to have authority [...] than have generally been realized” (Maitra 2012, 96).⁹⁰ She describes two such possible ways speakers can come to have subordinating authority where they once lacked it. First, through “derived positional authority,” where a speaker gains authority in virtue of inheriting a part of someone else’s basic positional authority; and second, via “licensing,” where a speaker’s audience grants her the authority she seeks through its actions or inactions (ibid., 109, 111). In the case of licensing—which will be my main focus here—Maitra’s aim is to fully disentangle authority from social position. In these cases, she argues that speakers can acquire authority over others without occupying or inheriting any *position* of authority whatsoever.

⁸⁹ While this is no doubt a vague category, the crucial distinguishing feature of *ordinary instances of hate speech* that is at issue in this paper lies in the fact that these speakers occupy no obviously relevant positions of authority. Moreover, it is worth remarking that these cases make up the vast majority of hate speech that is uttered.

⁹⁰ As will become clear later, I believe this is, overall, the correct strategy to take when answering the Authority Problem. For an account that does not take authority to be necessary for subordinating speech, see McGowan (2004; 2012). I discuss this alternative in more detail in the next chapter.

I believe that this account is deeply—and interestingly—mistaken, and in this chapter I raise some problems for the account of licensing speaker authority Maitra offers for subordinating speech. In particular, I take issue with what I see as its highly localized character, which effectively divorces the subordinating authority of ordinary hate speech from the broader normative context, including social features that I claim play essential roles in subordinating acts. For this reason, I worry the account is unable to properly distinguish between cases where a speaker does and doesn't have the authority to subordinate. Seeing this mistake points us towards a better account of speaker authority. For these reasons, I develop an alternative answer to the important question of how ordinary hate speakers can sometimes come to possess authority and offer a competing picture of how licensing functions in these cases. Crucially, as I defend it, the type of speaker authority that's required in cases of ordinary hate speech is still positional authority, though it need not be formal nor clearly defined. I call this type of authority *informal situational authority*.

I begin in section 2 by briefly explaining the standard argument for how speech can constitute subordination, which I first discussed in the beginning of Chapter One. I then describe an example of an ordinary instance of hate speech and make clear how it appears to fail in the face of the Authority Problem. Next, in section 3, I consider Maitra's account of licensing, which aims to answer this problem by demonstrating one way in which ordinary speakers can be invested with a kind of authority that is not tied to any social position. In section 4, I then show how her account seems to entail counterintuitive results, labeling cases as potentially subordinating where I believe the label shouldn't apply. I locate the source of this difficulty in the narrowly local picture Maitra defends, and explain the problematic implications this has for her account of the pragmatic structure of hate speech. In section 5, I argue that this problem

stems from Maitra's claim that licensing invests *people* with authority directly, and not through social positions. I believe that by appealing to social positions, even in licensing, we end up with a better picture of how some speakers can come to possess the necessary authority and subordinate with their speech. I suggest how this is possible by sketching an account of informal situational authority. Finally, I show how this approach also offers a better picture of what's required to challenge and defy these acts by those who may be seen to fall into the position of bystanders.

2. SUBORDINATING SPEECH AND THE AUTHORITY PROBLEM

Since the late 1980's, an influential movement in the philosophy of language has focused on the important intuition that one of the many acts speech can be used to do is to subordinate.⁹¹ A key development in this area has been the thought that some speech, such as racist hate speech, doesn't just lead to subordination, but is itself a subordinating act. That is, speech can subordinate. When probing what this meant in Chapter One, I focused on clarifying the distinction between causing and constituting. Here, I want to ask what is included in the concept of subordination.

In broadest terms, to subordinate another is to diminish his or her civil status. Getting more specific, we can identify three features of subordinating acts. First, they unfairly *rank* their targets as having inferior worth; second, they *legitimate* discriminatory behavior towards their targets; and third, they unjustly *deprive* their targets of some important powers (Langton 1993,

⁹¹ Catharine MacKinnon is acknowledged as providing the key first moves on this topic; see MacKinnon (1987). And while much of the early (and ongoing) work in this area has primarily addressed pornography as subordinating speech, many insights have been carried over to work on racist (and other forms of) hate speech.

303–04). Some explicit examples of subordination are laws that define enslaved people as the property of their masters, or women as the property of their husbands. These invoke a hierarchy of civil statuses and confer greater and lesser rights and powers to the different classes they define. And this doesn't need to be so explicit in the law, but can be embodied more subtly—though, often not that subtly—in the institutions, practices, and norms of a given society.

In what follows, I focus primarily on the act of ranking as inferior as a telling component of subordination, in part to more closely follow Maitra's approach to the problem (2012, 100). The question I aim to explore here is whether and in what ways ordinary hate speech can be seen to successfully issue rankings of inferior civil status to its targets, and thus possibly subordinate.

When it comes to how *speech* itself can constitute subordination, we saw earlier that the “paradigmatic” example is the apartheid-era South African legislator who, in saying “Blacks are no longer permitted to vote,” makes it the case that Black South Africans can no longer vote.⁹² According to Langton, this utterance—the very act of saying those words in those circumstances—constitutes subordination because it itself successfully *performs* the act of, among other things, unfairly ranking Black South Africans as inferior to white South Africans. That is, simply *in* saying what he does, the legislator successfully ranks his targets as inferior.⁹³

As others have pointed out, one obviously relevant reason the legislator's speech is able to subordinate in this way is because his authority is grounded in the particular formal position he currently occupies—namely, that of a legislator. A legislator has what we can call the *positional authority* to rank certain citizens as inferior to others in his jurisdiction, and this is all

⁹² Langton (1993, 302–03). As I note, I discussed this first in Chapter One, where I criticized its use as the ‘paradigmatic’ instance of subordinating speech.

⁹³ This distinction, of course, alludes to the Austinean distinction between perlocutionary and illocutionary acts. See Austin (1962, 101–09).

in virtue of being a legislator. Positional authority of this kind is what makes the legislator's ranking less like the simple expression of an opinion, and more like the rendering of a verdict.⁹⁴ Much in the same way that the umpire calling a strike makes it so, the legislator's positional authority is taken as a key reason why his speech act ranking Black South Africans as inferior is able to succeed, and thus for his speech to subordinate. More mundane cases of ranking share this important feature. For example, a teacher has the ability to rank her students in a class activity precisely because her position as a teacher gives her the authority to do so (Maitra 2012, 100). Try as they might, the students' parents cannot place their child at the top the class simply by saying it is so.

There are three further features of the legislator's authority beyond its mere positionality that appear relevant: First, there are *formal provisions* that place people in positions of authority; second, these are partly comprised by the particular *actions* of others; and third, the authority lasts for a *clearly demarcated period*.⁹⁵ In any case, it's useful to begin with the idea that it is positional authority of this kind that makes the legislator's ranking succeed.

With all this in mind, consider the following example of ordinary hate speech, drawn from Maitra:

An Arab woman is on a subway car crowded with people. An older white man walks up to her, and says, "F***in' terrorist, go home. We don't need your kind here." He continues speaking in this manner to the woman, who doesn't respond. He speaks loudly enough that everyone in the car hears his words clearly. All other conversations cease. Many of the passengers turn to look at the speaker, but no one interferes. (2012, 100–01)

Along with many others, I see this as a plausible example of a subordinating speech act. Yet, the speaker is not in any recognizable position comparable to the South African legislator; there are

⁹⁴ It is the legislator's positional authority that gives his speech the verdictive and exercitive force—to use the Austinean terms—that allow him to subordinate. See Langton (1993, 304–305).

⁹⁵ For more on these features, see Maitra (2012, 104).

no actions taken by anyone that formally confer authority on him for any clearly demarcated time period. He would therefore seem to lack the positional authority required to make successful rankings of others. This is the challenge of the Authority Problem. And so, while his speech *may be an attempt* to rank his target as inferior, absent the clear authority to do so, it seems this ranking must fail. To make things explicit: If occupying a formal position of authority—like that of a legislator—is necessary for one’s speech to succeed in *ranking* others, we’re then led to the conclusion that ordinary instances of hate speech like this cannot constitute subordination, quite simply because they don’t come from legislators, but ordinary speakers who lack the relevant positional authority. Avoiding this counter-intuitive conclusion and explaining how and why racist hate speech of this sort can indeed be subordinating is the goal of the remainder of this chapter.

3. MAITRA’S SOLUTION TO THE AUTHORITY PROBLEM

Maitra’s strategy in responding to the challenge presented by the Authority Problem is to disentangle authority from social position. Through her account of “licensing,” speakers can come to have authority over others without occupying any recognizable position of authority. This allows even ordinary hate speakers to gain the authority they need for their subordinating speech acts to succeed.

To explain this idea, she asks us to consider a group of friends trying to plan a hike together (Maitra 2012, 106). As no one expresses any strong preferences in the logistics of the trip, they fail to make much headway in actually planning the outing. After having enough of this, one friend, Andy, decides to take charge and make decisions. He assigns specific tasks to

each of the different members of the group. No one objects, the tasks are all completed, and the hike later takes place as Andy planned.

Maitra's claim here is that in this case Andy clearly comes to have the authority to assign tasks to the group, and that his "instructions, moreover, are authoritative speech" (ibid.). That is to say, his speech act constructs viable norms for his audience, and this can be seen by the group following his directions. However, as Maitra points out, Andy's authority is not in virtue of any clear social position. It's therefore, she claims, a type of speaker authority that *isn't* positional authority. Furthermore, it's only through the act of speaking itself that Andy comes to have any authority at all. That is, he has no authority prior to giving the instructions, but is granted this authority by the group only when no one objects to his orders and it becomes clear they plan on following through with the tasks he assigns—as Maitra notes, though, the boundaries here are essentially fuzzy. Finally, we may note that for speech to be licensed like this, it doesn't require "the licensors agree with the license in any substantive sense" but only that any reservations aren't made public (ibid., 107). So, some of Andy's friends may not be too pleased with his choices, and yet, so long as they keep these thoughts private, they are still expected to do as instructed—that is, the norms are still applicable. As such, disagreement—even strong disagreement—is compatible with an audience granting a speaker authority and thereby enabling her speech to construct new norms in social space. To summarize, then, here we seem to have a clear case of authority that's not positional, didn't require any formal granting of authority, results mainly from omissions rather than acts, and is perhaps open ended.

With this account of licensing in mind, Maitra claims we now have the tools to interpret the earlier example of ordinary hate speech as authoritative. As was previously said, the speaker can be thought to be issuing an attempted ranking that aims to mark his target as inferior. Given

what was said above, however, it's now possible to see how he succeeds with his ranking. In drawing the attention of all the other passengers in the subway car with his tirade, Maitra suggests we might think of the passengers all as participants in a single conversation, like the hike-planners (ibid., 115). And, by failing to voice any objection, we can now see the hate speaker's audience as licensing his speech, therefore granting him the authority he needs to make this ranking succeed.

As we just saw, the fact that the speaker occupies no formal position of authority prior to his utterance doesn't entail his audience cannot give him—even unwillingly—the authority his speech act needs. Like Andy's friends, the subway passengers' continued silence allows them to grant the speaker the authority he lacked prior to speaking. With their silence, the passengers refuse to challenge the hate speaker's ranking of the Arab woman as inferior, and this permits the ranking to add its content to the shared background of the conversation, making it successful.⁹⁶ And so, even ordinary instances of hate speech may in fact subordinate, since even ordinary hate speakers can be licensed in this way.⁹⁷

Now we should pause to note that this is not to claim this is what happens in every, or even most, cases of hate speech. Maitra's claim is more modest, saying only that there is no principled reason to outright reject the idea that ordinary instances of hate speech might sometimes subordinate in this way. It is a claim about sufficient conditions, not necessary ones. Yet, this is still a potentially significant result, as it would demonstrate that the authority to

⁹⁶ The idea of a conversational 'background' or 'common ground' appealed to here comes from Robert Stalnaker; see Stalnaker (2002).

⁹⁷ In a similar (though distinct) manner, McGowan argues that any contribution to a conversation invokes rules of accommodation, and that an ordinary instance of hate speech can constitute subordination in virtue of functioning as a "conversational exercitive," that is, an illocution that enacts permissibility conditions in a particular domain. Rae Langton, Sally Haslanger, and Luvel Anderson see this as one way to account for the Authority Problem. See McGowan (2004, 99–101); as well as Langton, Haslanger, and Anderson (2012, 759–60); and, for background, see Lewis (1979). I thank an anonymous reviewer for *Social Theory & Practice* for highlighting this connection for me.

subordinate need not always be tied to formal positional authority, like that of a legislator. In what follows I want to suggest that Maitra's account is half right. That is, I argue that Maitra is correct in saying that subordinating authority is not always formal positional authority, but incorrect in claiming that it is not positional authority, nonetheless. Indeed, I believe the type of authority that makes ordinary instances of hate speech subordinating is fundamentally positional, and intend to show this through a critique of licensing.

4. PROBLEMS WITH MAITRA'S SOLUTION

On Maitra's view of licensing, as we've seen, it's possible for subordinating speaker authority to emerge out of the facts of the local speech situation. By this, I mean facts about the content of what is said, and whether or not it is plausibly added to the conversational background, which can be determined on the basis of whether it is challenged or not by the speaker's audience. I believe this narrow focus on local features causes three related problems. First, this leads her view to provide counterintuitive results concerning when a speaker does and doesn't have the authority to successfully issue rankings and potentially subordinate, as it seems to ignore the broader social factors necessary for subordination. Second, and consequently, I believe this account has difficulty distinguishing between cases where an utterance has the performative force of a ranking, and ones where it functions more plausibly as an expressive, in that it is taken as primarily an expression of the speaker's wholly personal emotions rather than a claim about the world.⁹⁸ And lastly, as I will explain in the following section, Maitra's account

⁹⁸ The account of expressives I appeal to here—and explain more fully below—comes from Kukla (2014, 450–53). It is worth noting that Kukla is using 'expressives' in a non-Austinean way, and thus takes us away from the Austinean framework that Maitra employs. While this involves a slight change in terminology, I believe it is

has problems clarifying the domain of speaker authority, since it holds that licensing invests authority in a person independent of any social position.

Imagine a case similar to the subway example, except that instead of an Arab woman facing the brunt of an older white man, a well-dressed white man is accosted by a lone, unkempt, young male passenger. The young man calls the older passenger “a greedy f**k,” who “got rich by raping the planet,” and claims loudly that the world would be “better off without him” and his ilk. Just like in the original example, all other conversations cease, and no one, including the target, voices any objection to the statements being made.⁹⁹

Some—though not I—might be tempted to describe this modified case similarly to the original. That is, the speaker might be thought to be issuing a ranking that aims to mark his target as an inferior and undesirable. With sustained silence from the other passengers, his ranking succeeds. And since unfairly ranking as inferior is a feature of subordination, we might conclude that this younger man subordinates his target in saying what he does. Yet, I believe it is very unlikely that this example depicts a case of subordination, and figuring out why can help us understand the type of authority at play in ordinary instances of hate speech.

The first and most obvious way to avoid this counterintuitive result would simply be to point out that what’s gone wrong here is that this ranking is not an *unfair* or *unjust* ranking, and therefore not subordinating like the original case.¹⁰⁰ We might say, then, the ranking succeeds,

ultimately fruitful, in part because it allows us to examine the issue of uptake more broadly. I thank an anonymous reviewer for *Social Theory & Practice* for encouraging me to clarify my shift in terminology here.

⁹⁹ Much like the original subway case, this one is likely not unfamiliar to anyone who has lived in a larger city for some time, and I chose it in part because of its familiarity. As such it should be thought of as taking place within an ‘everyday’ context, rather than, say, at a political rally, which would be relevantly different.

¹⁰⁰ Of course, intuitions likely conflict on this point, and many, including an anonymous reviewer for *Social Theory & Practice*, would see this as an unfair or unjust act—though perhaps not a subordinating one. In any case, my point is simply that there is a lack of symmetry between the two subway cases under consideration, and we need to investigate their differences more closely.

but it's not subordinating because its content is not morally objectionable. This distinction between fair and unfair, or just and unjust, acts of ranking helps to separate the troubling case of the South African legislator from the more mundane case of ranking one athlete as faster than the others (Langton 1993, 304). So, it must be registered that this second subway case does not involve anything near the type of *unjust* ranking found in the initial example. I take it as given that the two subway cases differ dramatically in their content, in that one is deeply unjust while the other is arguably anti-oppressive. This has clear implications for the subordinating force of the one in contrast to the other.

But I won't pursue this strategy, in part because it makes 'unjust' do a lot of the crucial work, and I'm somewhat skeptical it's up to the task. What I want to point out, moreover, is that this is not *all* that is different in these two cases—the explanation is not that simple, though it seems to be the (only) one that Maitra's account permits. Even once we acknowledge this distinction in the content of the utterances in the two subway cases, the young man's utterance is still clearly quite different from the completely innocuous form of ranking done at the end of a 100m dash or after an ice dancing routine. First, it is not obviously the mere registering of an empirical fact (as the 100m dash is), but rather involves the projection of specifically ethical values (and so, unlike ice dancing). And second, it issues a ranking not about athletic or artistic ability, but about one's standing in a social hierarchy. Arguably, any departure from equality here is *prima facie* unjust.¹⁰¹ This is not to say it comes anywhere near rivaling the type of injustice at issue in our first case, it is only to note that I find this route—that while both cases describe the successful issuing of a ranking, only the first is an *unfair* ranking and that is where

¹⁰¹ One broad camp of egalitarians—Relational Egalitarians—can be thought to oppose any social system that endorses *any* hierarchy in the area of social standing; see Anderson (2012, 43–44).

the difference lies—too quick and too blunt. Furthermore, I just don't see these two cases as being even *this* similar to each other, even aside from their differences in content.

To this end, I think it's best to resist the notion that the young man's attempt at ranking his target as inferior, whether fair or unfair, succeeds. One way to do this is to see why the passengers may not take the young man to be attempting a ranking at all, even if this is exactly what he intends to be doing. That is, whether some utterance has the performative force of an assertion or of an expressive is sometimes not up to the speaker, but rather partially constituted by the uptake it receives (see Kukla and Lance 2009). As Rebecca Kukla argues, the difference between these two speech-acts—assertions and expressives—lies, in part, in how assertions make truth claims “about *how the world is*, and it either is or isn't that way for *everyone*” (Kukla 2014, 450, emphasis in original). In doing so, assertions seek “uptake in the form of agreement or rational challenge” (ibid., 451). Expressives, on the other hand, may be thought of as mere expressions of one's feelings, and therefore as “‘wholly personal,’ non-truth-bearing speech acts” (ibid.).¹⁰² This difference shows up in the uptake expressives receive. As Kukla points out, “it doesn't even make sense for me to ask whether I agree or disagree with [your expressive], or whether it reflects the world correctly. [...] It is merely an expression of feelings, disconnected from rational discourse” (ibid.). I want to suggest that the young man's utterance in our second example is best understood as functioning as an expressive in his situation. What he perhaps intends to be a statement about the world, the passengers might understandably take as an outburst of anger, and respond accordingly.¹⁰³ Their silence, therefore, might not be indicative of

¹⁰² Kukla takes the phrase “wholly personal” from Naomi Scheman; see Scheman (1993, 24–25).

¹⁰³ Here it is worth acknowledging that rankings and assertions are not identical acts. However, I agree with Maitra when she claims that the two are importantly related, and that “a speaker can perform a ranking just *by* asserting something;” see Maitra (2012, 112).

their granting him the licensed authority to issue a successful ranking, but, instead, tolerance for his anger, perhaps.

But I don't want to say the same thing about the ordinary hate speaker in our first case. Indeed, my basic claim is that these two cases share very little in common at the level of pragmatics. That is, beyond the important differences in the content of our two cases, it is the differences in their pragmatic structure that provides crucial insight into the subordinating authority of ordinary hate speech.

As I see it, even if the hate speaker sees himself as merely voicing his own 'wholly personal' anger, it's important to see something else, something very close to Maitra's description of the case, going on. No matter his intentions, in speaking as he does—and being who he is—I take it that he performs a subordinating act. We might see this as the flip side of a phenomena Kukla terms “discursive injustice.” Cases of discursive injustice occur when a speaker who has the standard entitlements to perform a certain speech act goes about performing this act in all conventionally appropriate ways, and yet *because of her gender*, the uptake she receives constitutes her speech act as a different kind of act than its input would conventionally dictate (2014, 444–45).¹⁰⁴ An utterance that would standardly be an imperative might become a request, simply because it comes from a woman in a traditionally male-dominated environment. And so similarly, despite the fact that ordinary speakers are not normally in a position to construct norms for others with what can be seen as potential rankings, the social position of *some* ordinary speakers, like the ordinary hate speaker in our original subway case, empowers their speech with unique performative force.

¹⁰⁴ Kukla notes that this phenomena is not restricted to instances where gender is the salient feature; discursive injustice also occurs where race, ability, orientation and other forms of membership in disadvantaged groups are at the fore.

I take it that something along these lines helps to explain the differences in the pragmatic structure of the two subway cases, as I highly doubt the ranking of the man in the second subway case is—aside from content—identical to that of the Arab woman. Yet, if the authority to rank others can be an emergent fact of the local speech situation, and if the deployment of rankings, given the right response, add their content to the shared background, something close to identity in pragmatic structure seems to be a plausible consequence of Maitra’s view. As Maitra puts it,

[w]e might say the ‘essential aim’ of assertions is to make a difference of a particular kind to the shared background for a conversation, namely, to add its content to the shared background, and so to eliminate all possibilities incompatible with that content. ... [And] we can say that a ranking, just like an assertion, seeks to add its content to the shared background of the conversation. If none of the participants to the conversation objects, the content may be added. That means that the background must be updated to eliminate all possibilities incompatible with that content. (Maitra 2012, 112)

My worry is that what distinguishes the second subway case from our original example are features that Maitra’s account of licensing has difficulty explaining.

This localized understanding of authority is clearest when Maitra says that in the case of licensing, the scope of authority “extends only to the parties to the conversation, and no farther” (2012, 117).¹⁰⁵ She goes on to explain how the overlapping feature of multiple instances of hate speech “reinforce the subordination that *each instance separately contributes*” (ibid. 117–18, emphasis added). On this picture, then, the main difference between the two cases is in how often each occurs; their internal features are, in many key ways, the same.

More importantly, I worry that an account of licensing of this sort runs the risk of ignoring the fact that hate speakers are never starting from scratch.¹⁰⁶ It’s simply not the case that

¹⁰⁵ To be clear, Maitra is here describing the “jurisdictional” question of authority for cases of licensing.

¹⁰⁶ This is similar to the argument I made in Chapter One about Langton’s use of the South African legislator’s speech as a paradigm example of subordinating speech. Like Maitra, she over-inflates the role of an isolated speech act from a single speaker. I take these to be distinct, though obviously related criticism.

the speakers in our two subway examples above are making anything close to similar moves with their speech acts. What's needed is an account of how ordinary instances of hate speech draw on, exploit, and give voice to pre-existing oppressive social norms that they then reshape and deploy in their utterances. Indeed, I take it that it's partly because ordinary hate speech is so ordinary, in the sense of not uncommon or familiar, that it is able to have the authority it does, and this element of familiarity needs to be recognized in our account of subordinating authority.

5. INFORMAL SITUATIONAL AUTHORITY

To begin to build an alternative account, I want to appeal to Lynne Tirrell and her explanation of how our social interpretations and practices distribute “situated power” to particular individuals (1993, 14). This situated power is distributed, in part, through what Tirrell calls the “default interpretations” that one’s community enables.¹⁰⁷ And while Tirrell’s focus in that paper is the default interpretation that a man is a full agent, and a woman is “less so” (ibid., 14), to this we can add racist interpretations that depict people with darker skin as suspicious, as more likely than whites to commit crime, etc. When these interpretations are backed up by a community—such as by norms, laws, and institutions that enable and encourage the police to stop, search, and ultimately detain and incarcerate anyone they deem ‘suspicious’—they empower some individuals more than others, and allow them to perform certain acts that are unavailable to the disempowered. Tirrell claims, and I agree, that “this conception of situated power helps to bridge the gap between social practices and individual experience” (ibid.).

Returning to the two subway cases: the ranking of an Arab person as inferior and

¹⁰⁷ Tirrell credits this concept of situated power to Tom Wartenberg; see Wartenberg (1988).

unwanted because of a possible connection to terrorism is, while deeply unjust, an association that is sadly backed up by the broader community.¹⁰⁸ The hate speaker is best thought of, then, as giving voice to this pre-existing racist interpretation. He is functioning more like a mouthpiece for this interpretation, and less like a self-originating source of new norms. But crucially, it is also because of his status as a white man that his assertions are more likely to carry real weight in his community; indeed, they are more likely to be taken as assertions with truth-content rather than wholly personal expressives. These features fail to carry over to the modified example. But notice how neither—not the well-known racist association of his rant, nor the excessive standing given to white men in some societies—is established in the local speech situation; rather they are part of the background social norms that inform the local situation from the outside. Yet, an account of licensing like Maitra’s is isolated from these vital aspects of subordinating hate speech.

These potential oversights in Maitra’s account, I now want to suggest, ultimately stem from her claim that licensing invests *people* with authority directly, not through social positions.¹⁰⁹ For example, in the hiker case, Maitra says that “licensing invests Andy with authority. It does not invest authority in whoever occupies some given position” (2012, 114). My suggestion is, rather, that this is exactly what licensing does; the authority that speakers come to have through licensing is in virtue of *creating* and *moving into* relevant positions of authority, even if highly local and ad hoc ones. Andy’s authority emerges from his coming to occupy the position of *hike organizer*—a position that only exists because he creates it and moves himself

¹⁰⁸ What it takes for an interpretation to be backed up by the broader community is an important question that is, unfortunately, too big for me to answer here.

¹⁰⁹ It is clear that Maitra takes this to be an important upshot of her paper, as she says that “having authority isn’t the same as occupying a position of authority” (Maitra 2012, 114).

into through his speech. I suggest we call this type of authority *informal situational authority*.

To get clearer on this idea, we might think of the initial instructions that Andy issues as functioning as a *meta-call*. As Mark Lance and Rebecca Kukla define it, a meta-call is “one in which the primary function of the call is [...] to restructure normative relationships and possibilities for making first-order calls” (2013, 473). In the hiking case considered above, Andy is seeking the entitlement to issue commands to his friends—something he does not have standing to do at first. Yet, in uttering commands that require some authority to be binding, Andy can be thought to be both asking for this authority and marshalling it at the same time. His words, then, “do double duty” (ibid., 474–75). While his commands aim at his friends’ compliance, more fundamentally his utterances seek to alter his status relative to his audience, putting him in a position to determine the norms that bind them all.¹¹⁰

Moreover, I believe this framework helps to resolve a possible ambiguity in Maitra’s account. While she explicitly says licensing invests people with authority, she also says it is the person’s *speech* that is licensed, and what’s more, that they are only licensed within a restricted domain, that is “someone is licensed to perform certain actions that affect certain others” (2012, 107n).¹¹¹ This explains why Andy can only order his friends to do things relating to their hike. If he tried to go beyond this domain, say, and ordered his friend Tomas to breakup with his boyfriend, it’s very unlikely this would be taken as an authoritative speech act. Yet, on the view that licensed authority is both attached to the *person* and open-ended, it’s hard to see what rules this out so obviously. You could of course say this order is unlikely to go unchallenged and so

¹¹⁰ We might then think of his first utterance as functioning primarily as an entreaty. An entreaty, for Lance and Kukla, is a meta-call that “calls [on] someone to grant to the caller an entitlement to make certain kinds of claims that the called is not yet in a position to make” (2013, 474).

¹¹¹ Moreover, at another point she says because of “licensing, the speakers are in a *position* to perform their intended acts” (Maitra 2012, 113).

would not survive the licensing process. But it seems unnecessary to appeal to a failure of licensing here when the order looks, on the face of it, to lack any pretense of authority given the established context. Besides, I think it's best to resist this more piecemeal approach to licensing, since, as Maitra suggests, it seems the authority in question is temporally extended in a fuzzy way. As I see it, the clearer explanation is because Andy's authority extends only as far as is determined by the social position he now occupies. The normative territory he enacted for himself and his audience is bounded in this more structured way. Even licensed authority is best understood as positional authority. And this makes sense since authority is a relation between two or more parties, where one party is granted more power than the other. Seeing speaker authority as positional makes this relation vivid. Yet, as Maitra does well to illustrate, this authority need not be tied to formally recognized positions, like that of a legislator. This, however, makes them much harder to spot and easily missed.

How does this help with the two subway cases? I believe an analysis along these lines provides an entry point for the broader social facts that a pure licensing account misses. This is because the social roles that come with certain amounts of positional authority are, in many cases, closely related to the default interpretations backed up by a community. We see this when a white man's opinions are more likely to be taken up as the concerns of the community at large, while a Black woman's are often taken as indexed and fixed to her racial community.¹¹² This is partly because, as Tirrell notes, “[f]or the oppressed, identity factors tend to limit uptake”

¹¹² Caroline West discusses this imbalance in terms of a “consideration failure,” where the opinions of marginalized persons do not receive the same consideration in the dominant society as would dominant speakers. See West (2012, 244–45). West's account is partly built off the work of Charles Lawrence III, who, on this subject, notes that: “Racist speech [...] distorts the marketplace of ideas by muting or devaluing the speech of Blacks and other despised minorities. Regardless of intrinsic value, their words and ideas become less saleable in the marketplace of ideas. An idea that would be embraced by large numbers of individuals if it were offered by a white individual will be rejected or given less credence if its author belongs to a group demeaned and stigmatized by racist beliefs” (1993, 78–79).

(2018b, 26). In this way, the social position of a white man seems to come with more authority to dictate what is and isn't acceptable in public space, because in a racist society, it is taken as *his* space more than it is *her* space.¹¹³ This crudely simplifies matters quite a bit, of course, but I believe it gets us closer to understanding how ordinary instances of hate speech like the subway case can constitute subordination. In the original subway case, the man's speech can be considered authoritative not simply because the passengers' silence licenses *some person* to rank however he pleases, but because in (a) saying what he does, and (b) being who he is, he more easily moves into a role with the informal situational authority to say what is and isn't accepted in the broader community. As an *arbiter of community standards*, say, his ranking may be thought to have originated from beyond the local speech situation, and thus carry real weight. It is therefore in taking up and giving voice to broader social views that an essential aspect of the ordinary hate speaker's authority lies.

I take it that a similar move is not—or at least less—available to the young man in the second case as his statement is less likely to function as a successful entreaty to an authoritative social position, and more likely taken as the angry tirade of a dissatisfied youth. And in the role of *dissatisfied youth*, he has no authority to determine the social statuses of others—save for maybe who is cool or not. In this way we can see how the licensed speaker authority that Maitra identifies is not independent of positional authority, but rather one way among others people may move into social positions of authority.

However, even though I believe we have good reason to describe cases like ordinary hate

¹¹³ As Tirrell puts it: “There are no universally neutral entrances, for the salience of identity factors is always in play and highly dependent on social context. Identity factors, historically entrenched in practices, often influence whether someone will even try to enter a game, whether others will try to keep her out, and how such exclusions would be achieved” (2018b, 25).

speech as cases of positional authority, they aren't so simply described this way like the case of the South African legislator may be. That is, we must also aim to capture the dynamic process that Maitra describes occurring in licensing. This is important both to fully understand the mechanisms through which authoritative speech is enabled, and also to provide helpful guidance for how such speech acts should be resisted.

One of the substantial virtues of Maitra's account of how ordinary hate speech gets licensed is how it holds the licensing audience responsible. If it is because of their silence and refusal to challenge the hate speaker's speech that his ranking succeeds and therefore subordinates its target, then there is a clear moral obligation for the audience to speak up and block this attempted ranking. Silence is not a morally neutral response, and private disagreement does nothing to prevent subordination. In this way, she highlights the complicity that can be attributed to the audience of such an instance of ordinary hate speech (2012, 116–17).

However, once we see how this authority is, *pace* Maitra, still positional authority, we're led to a different understanding of what it means to challenge this act. To begin, I believe that Maitra underestimates what it takes to resist subordinating acts. If one's moral complicity in subordination is constituted by their participation in licensing a speaker's utterance, then it seems that one can discharge this responsibility rather easily. Recall, the feature that allows an audience to license a speaker's subordinating ranking despite strongly held reservations is the fact that so long as these reservations remain private, they fail to challenge the utterance's content being added to the conversational background. However, consider the situation had a single passenger in the original subway case spoken up, saying "That's not very nice" to the ordinary hate speaker. To be sure, this is, given the context, a potentially brave thing to do and is much better than remaining silent. At the same time, though, I fail to see why this should render all of the

hate speaker's subordinating authority null. Indeed, it seems plausible to me that even in a case where a passenger challenges the hate speaker's words in this way, this act still constitutes subordination.¹¹⁴

I believe the reason why this only goes so far as a response, and why it fails to challenge in the right way, is because it is directed at the hate speaker's utterance at level of content, and not the speaker's authority more directly. In their analysis of the pragmatic structure of how a speech act may be met with defiance, Lance and Kukla suggest that "defiances can be divided into responses that serve to challenge the speaker's entitlement to the original speech act and those that do not" (2013, 470). On the one hand, the "That's not very nice" response clearly challenges and resists the original speech act—it does not allow its unfair ranking to be added to the conversational background unobstructed. On the other hand, this response does not seem to challenge the speaker's entitlement or standing to make a call of this sort. That is, what's being challenged is the precise ranking that's being made, and not the fact that this speaker claims to be in a position to issue a ranking at all. As such, it may be thought that by challenging the utterance in this way this response still "acknowledge[s] the force of the norm that [the passenger is] violating" (*ibid.*, 471). That is, there is a disagreement about the content at issue, rather than the hate speaker's standing to issue content of this sort. This would be similar to a tennis player's disagreement with the umpire's decision to call "Out!" when she believes it was 'in,' rather than if she wanted to challenge the umpire's entitlement to even make a call at all, perhaps because she questions the umpire's training, attention, neutrality, and so on.

In order to fully challenge the hate speaker's utterance, what's needed is an act that

¹¹⁴ I explore this potential problem for accounts licensing and accommodation more fully in Chapters Three and Four.

resists the speaker's attempt to step into an unjust position of authority over others, or one that challenges the authority of the utterance directly, rather than simply challenging what he does *with* that authority (ibid.).¹¹⁵ Unfortunately, there cannot be any universal principles on how this may be accomplished in the varied contexts where ordinary instances of hate speech occur. In general though, acts that aim to reposition the ordinary hate speaker as someone who clearly *lacks* authority over his audience, either by blocking his attempt to raise his own status or by making one's solidarity with the target of his speech clear, can help resist the subordinating force of his utterance. Moreover, once we understand that the subordinating force of ordinary hate speakers' words come from their ability to situate themselves—and not just their words—in a distinct normative relation to their audience, we have a clearer target at which our defiant acts must aim.

6. CONCLUSION

In this chapter I criticized Maitra's account of licensing for focusing solely on the local facts of the speech situation at the expense of ignoring the broader social facts that are essential to understanding speaker authority in cases of ordinary hate speech. I argued that this misstep means that her account fails to distinguish between cases where a speaker does and does not successfully rank/subordinate his target. This is the case, furthermore, because Maitra's account has difficulty fully capturing the pragmatic structure of hate speech.

With Maitra, I share the intuition that ordinary instances of hate speech constitute

¹¹⁵ More dramatically, there is what Lance and Kukla dub an "activist speech act," that is, "one that does not merely draw upon but in fact functions to subvert or reconfigure the normative context in which it operates" (2013, 473). One possibility is that activist acts of this type may be supererogatory in some situations.

subordination, and also that subordination generally requires authority. Yet, I argued that authority of this sort requires occupying a particular position of authority. I suggested that the licensed speaker authority that Maitra identifies is not independent of positional authority, but rather one way among others people may move into social positions of authority. These positions can be ad hoc and informal, only emerging when particular utterances are attempted and given particular uptake but are still needed to help explain the domain of authority better than a pure licensing account. As I see them, cases of ordinary hate speech are cases where a speaker seeks *informal positional authority* via a speech act that functions both as meta-call aiming to give the speaker a novel authoritative standing and a ranking that enforces an unjust hierarchal civil status. Moreover, I claimed that positional authority of the kind that concerns subordinating hate speech often hides in plain sight, in the default interpretations backed up by communities.

Who does or doesn't have the authority to subordinate is not an easy question to answer. But it is not one that can be answered simply by looking at who holds the formal offices of authority. And it is equally not one that can be answered by looking solely at the local speech situation. Speech, and especially authoritative speech, must be seen to draw upon the social norms of the broader community. One way this might occur is by restructuring the informal positional roles that lay in waiting for speakers to take them up. And as with most positions of authority, they are often more easily accessed by some members of the community than others.

CHAPTER THREE: PLURALISM IN SUBORDINATING SPEECH ACTS

1. INTRODUCTION

So far in this dissertation, I have sought to add depth to the claim that some speech acts are oppressive and harmful, accounting for the importance of both causal and constitutive interpretations. In doing so, I have also defended the claim that some of these speech acts require the speaker to possess a type of authority to accomplish this subordination successfully but argued that this authority can be much more ad hoc and informal than is usually recognized. This chapter continues to explore how speakers engage in oppressive speech and addresses the diversity of entitlements that make different subordinating speech acts possible. So, I ask what types of authority enable subordinating speech in addition to the type of licensed informal situational authority of subway hate speaker considered in Chapter Two.

Below, I sketch a rough, non-exhaustive taxonomy of subordinating speech acts and their entitlement conditions. It is non-exhaustive because I do not presume to cover the whole terrain of subordinating speech in a single chapter. The language of bigotry is always evolving, and any analysis will always be but a snapshot. The rise of the internet and social media has led to many new developments in this terrain, including the use of spamming, mobbing, and doxing as weapon in the hate speaker's arsenal. And so, I cannot hope to include everything here—though I do explore some of these notable aspects of online speech in more detail in the following chapter.

And it is also rough, moreover, because I do not believe the distinctions to be drawn below admit of rigid boundaries. Rather this is a mixed bag of speech acts. Consider how some

target an entire group (e.g., all Muslims), while others target a specific individual belonging to several groups (e.g., a particular, identifiable, hijab-wearing Muslim American woman on the political left). But even in the latter case it's clear that other Muslim women—and other non-women Muslims, and other non-Muslim women, etc.—may be harmed as well. Identifying the direct target of a subordinating speech act does not exhaust those whom it harms, but I believe it is an important first step in understanding what speech act is being performed.

In other cases, it is the speaker whose social position or identity is relevant to its status as a subordinating speech act. For example, when it's the president uttering what is arguably hate speech, that is significant in a somewhat unique way. At other times, it is the speaker's *anonymity* that lends their speech act subordinating power (e.g., a 'White Power' flyer posted on lamppost). In-between cases are prevalent here as well, where the speaker is not anonymous, but also does not occupy a unique position of authority (e.g., a group of male colleagues sharing sexist 'jokes' over company email).¹¹⁶

Despite these many differences, the speech acts I examine all share the feature of being a tool of subordination, or at least the potential to function as such. They share the potential to cause or constitute the types of harms discussed in Chapter One. This is therefore their essential pragmatic effect. This is what they *accomplish*, and this is what binds these speech acts together, and unifies them under a single concept—subordinating speech. Yet, as I will show below, there are myriad paths open for speakers to achieve this effect, and these differences form part of the explanation of how these constitute difference acts. So, while the boundaries are unquestionably rough, I claim that acknowledging differences like these opens useful avenues of analysis.

¹¹⁶ Or, for a recent real-life case, a group of United States Custom and Border Patrol (CBP) agents were uncovered to be sharing hateful content in a secret Facebook group.

I believe that reflecting on the diverse sorts of acts that together constitute the category of subordinating speech is useful for a few reasons. First, this chapter offers resistance to the idea that subordinating speech acts all collapse into one category—which we might call the ‘subordinative.’¹¹⁷ While no theorist explicitly argues for this view (as far as I know), it is often implied in their discussions on the subject. For instance, Mary Kate McGowan rejects the necessity of speaker authority and argues for a view that, in barest form, holds that “speech must be exercitive in order for it to be an act of oppression” (2009, 392). Similarly, Rae Langton advances what I will call a “monolithic” account of speaker authority for analyzing subordinating speech. Though she’s helpfully added various elements of nuance with recent modifications, her account still operates with the fundamental assumption that speaker authority should be modeled on the paradigm case of the legislator. And so, unlike McGowan who deflates speaker authority altogether such that it appears everywhere, Langton sees speaker authority as always indexed to some specific person. While I will have more to say about each of these views below in section 4, here I only wish to note that they seem to imply that all (or most) subordinating speech act function in roughly the same way. For my part, I am skeptical of these reductionist moves, and aim to avoid both extremes represented by these accounts: overly fetishizing the speaker-authority-as-legislator model of Langton, or overcompensating with the speaker-authority-is-everywhere (and so plays little or no role) model of McGowan. Both, I argue, fail to properly account for the speaker’s role in subordination.¹¹⁸

Second, there is the plain fact that we must understand these acts of subordination in order to combat them. As I understand it, the ultimate goal of an analysis of subordinating speech

¹¹⁷ John Michael Ramsey also argues against such an approach, though his specific topic and arguments differ significantly from mine; see (Ramsey 2013, 76–88).

¹¹⁸ Conversations with Matthew Shields were helpful in clarifying my views here.

ought to be to fully capture the complexity of this phenomenon, primarily as it is experienced by those who most often are its targets. A step I take towards that goal is to take stock of what has been left unexamined or under-examined in the dominant approaches.¹¹⁹ This chapter argues that the variety among subordinating speech acts is one such area left under-examined, especially by theorists following the narrow Austinean path that privileges the illocutionary act, represented below chiefly by Langton and McGowan. By identifying distinct subordinating speech acts—and different entitlements speakers draw on to accomplish these acts—we are in a better position to challenge and undo the harms of oppressive speech. Once we appreciate the different manners in which speech and language oppress, we understand better what we’re up against.

Finally, one addition I make here that will be important going forward is to identify a type entitlement for subordinating speech acts that is distinctively *collective*, in that it draws on a status provided by a group of speakers. That is, these speech acts attain subordinating force by being a part of a broader group-practice, where the individual speaker and their utterance would have little significance apart from their participation in this practice. The point I aim to make here is distinct from—though certainly related to—the general one about oppression itself being a systemic set of practices. I drew on that feature to highlight the deficiencies of illocution-focused accounts in Chapters One and Two, in that they tend to isolate speech acts from the broader practices within which they are a part. Here I point towards an entirely discursive version of this problem by considering speech acts that derive their distinct oppressive authority from being a part of a group of subordinating speakers. My aim is to show how the entitlement to

¹¹⁹ This is not to say that philosophers have left this topic untouched. Alexander Brown, for example, has recently given a thorough overview of the different types of hate speech and the corresponding hate speech laws that various jurisdictions have enacted or considered (Brown 2015). However, his project was mainly couched in legal concepts, while I consider this topic using the lens of social philosophy of language. Specifically, I use the tools of speech act theory and pragmatics to examine differences at the level of entitlement conditions and pragmatic force for the language-based mechanisms of oppression.

subordinate is not always located in individual speakers, but often in their participation in harmful group practices of subordinating speech. Chapters Four and Five will then continue this analysis through two revealing cases: online harassment and protest, respectively.

1.1. Methodological Foundations

In the introduction of the dissertation, I briefly described the role variance plays in the expanding philosophical literature on slurs. In short, the sheer variety of slurs themselves has led many theorists to consider differences across usages to be an important feature of slurs. This fact of variance can be considered a foundational piece of data that theories of slurs must account for if they are to be successful. Yet, in much of the literature on subordinating speech beyond slurs, the importance of variety plays no similar foundational role, and is often ignored. But, as I see it, the importance of accounting for variance with regard to slur use provides a *prima facie* case for a similar importance for subordinating speech acts beyond slurs.¹²⁰

And, to be sure, many theorists writing on subordinating speech do acknowledge variety in one form or another in their analysis. I am not the only writer making this point, by any means. Lynne Tirrell, for instance, has in recent work developed a conception of toxic speech, exploiting the metaphor of toxins and epidemiology in a fairly literal way. In doing so, she has done well to highlight the fact of variability. This is noted mainly in terms of harm in her

¹²⁰ It is somewhat odd that this stark difference could exist between philosophical approaches to slurs on the one hand, and philosophical approaches to subordinating speech on the other. However, it is a predictable result of the way how each subfield has grown largely independently of the other. Throughout this dissertation my focus remains for the most part on the subordinating speech subfield, which has its origins in Langton's speech act approach to pornography and later hate speech. The literature on slurs differs in part because of the importance semantic analyses have played in its overall development, which took this subfield in a different direction in general. Despite these contingent differences, however, theorists of subordinating speech surely have much to benefit from engagement with theories of slurs, and vice versa. Taking stock of the importance of variety of oppressive utterances, I contend, is one crucial lesson the analysis of slurs can teach us all.

account. “We should not assume,” she cautions, “a one-size-fits-all conception of harm; variability is the norm” (Tirrell 2017, 149). And, because of the epidemiological model that she is working with, she takes these differences in harms to be analyzable in terms of differences in ‘exposure.’¹²¹ In exploring this idea, Tirrell writes:

Some speech acts are very potent, delivering harmful messages that undermine decent functioning nearly immediately, and some of these cases have long-lasting effects. Some speech acts with a seemingly milder level of disrespect (for example) gain their power from repetition, so that the person might not be much damaged by any single event, but mightily damaged by the deluge of these across her lifetime. Route of exposure would likely map onto one-on-one speech vs. public speech, written vs. oral, known speaker or anonymous. Dosage would include both quantity and frequency, so would overlap with potency to some extent. Susceptibility factors would help guide our understanding why uptake varies. (2017, 147)

As this shows, the toxic speech model encourages an analysis where speech acts are understood to have a range of plausible effects, which are partly but not wholly fixed by the type of act it is. Different ‘doses’ and ‘routes of exposure’ produce different consequences in different populations. Tirrell uses an inferentialist approach to map these relations, with language entry moves, language-language moves, and language exit as the three (broad) types of speech acts used to categorize any instance.¹²²

This is a useful approach, and one I partly follow below.¹²³ But my sketch of the different types of subordinating speech acts differs in a few respects. First, I dispense with the epidemiological analogy, and aim to speak literally about the harms of subordinating speech. The

¹²¹ The importance of taking the idea of toxicity onboard is evident when Tirrell writes the following: “Medically, toxicity’s harm admits of degrees, and can be acute or chronic, depending on the number and potency of exposures. A poison’s impact and degree of hazard depends on its potency, route, dosage, and susceptibility or uptake. Susceptibility depends on the condition of the subject. Once exposed, some effects may last a lifetime, while others may be overcome” (Tirrell 2017, 147).

¹²² This framework builds upon systems developed by Sellars (1954); Lewis (1979); and Brandom (1998).

¹²³ Other work that was influential in my thinking in what follows includes Sbisà (2002; 2018); Berdini (2013); Hanrahan and Antony (2005); as well as Butler (1997). For an illuminating critique of Butler, see Schwartzman (2002).

metaphor to toxicity is a helpful entry point to capture the seriousness and the diversity of subordinating speech, but as a metaphor it opens the door to critics who might question the dis-analogies between literal toxins and speech. Second, I eschew the Sellarsian language of language-entry and language-exit moves. While helpful in certain contexts, here it has the troublesome implication that it runs the risk of implying that language and actions are two distinct modes of interaction; that the world of speech and the ‘real’ world are not the same one. I believe this implication ought to be avoided when discussing subordinating speech, as it risks playing into the ‘just words’ criticism. To be clear, though, this is by no means meant as a decisive objection. Rather, I am just laying my cards (preferences) on the table.

Instead, I opt for thinking of speech acts in terms of inputs and outputs. In broad outline, this follows the account developed by Rebecca Kukla and Mark Lance, where they maintain that “speech acts can be productively analyzed in terms of the normative statuses that enable them and the normative changes they effect through their performative structure” (2009, 13). This leads to a functional analysis of speech acts in terms of the normative statuses that entitle their performances, and the changes their performance tends to bring about. As they say:

If speech acts function to bring about changes in normative status, then they take normative statuses as inputs and produce them as outputs. Specifically, we can distinguish between the norms governing the proper production of a speech act, which give rise to statuses that entitle its performance, and the changes in normative status that their proper production strives to make. (Kukla and Lance 2009, 14)

In a similar spirit, I analyze subordinating speech acts according to what type of speaker entitlements they presume, and what type of uptake they (typically) generate. As I conceive of them subordinating speech acts all have the common goal of maintaining hierarchies of oppression. This is what they, as speech acts, strive to do, regardless of whether their speaker

shares this intention.¹²⁴ ‘Innocent’ microaggressions, racist ‘jokes,’ and ‘just trolling’ misogyny are all examples of subordinating speech acts that entrench oppression without this being the putative intent of the speaker.

Note, too, how the type of speech act is not always clearly indicated by its grammatical form. Questions, for instance, can have the subordinating pragmatic force of verdictives in the appropriate circumstances. Consider a recent tweet from the magazine *The Economist* asking, “[s]hould transgender people be sterilized before they are recognized?” (Levine 2019). This speech act combines a lack of (conscious) hostile intention with the grammatical form of a question, but nonetheless subordinates. Critics objected—rightly—to the question being presented so bluntly, absent any mention of its context regarding a decision from the Japanese Supreme Court. In doing so, *The Economist* implicitly legitimized the premise of the question, which casts doubt on the equal moral status of transgender people. This demonstrates how grammatical form is only a rough guide to subordinating force, as speech acts of likely any form can have the function of subordinating.

But aside from simply entrenching subordination, these speech acts have other notable features in their output worthy of differentiation. Rather than looking at speaker intentions—as if that were even really possible—or grammatical form, I examine the output of subordinating speech acts in terms of three features: to whom is it primarily addressed; whom does it target for subordination; and what normative statuses it assigns, which is informed in part by the first two features, along with the specific entitlements it marshals.

¹²⁴ Here I follow Kukla and Lance, who note: “The function of a speech act should not be confused with either the intention of the speaker in uttering it or the standard use of that string of words in the community. Of course [...] there has to be at least a defeasible concordance between function, intention, and standard use, and it is patterns of intention and use that serve to institute the contentful pragmatic structure of a language in the first place. But [...] we reject any analytic identities between function, intention, and convention, even while acknowledging that there is a constitutive, defeasible connection between these things.” (2009, 13–14)

On the input side, authority, as I understand it, is a particularly important type of speaker entitlement, but it is not the only kind. The discussion of the previous chapter may then be seen as an analysis of *one type* of entitlement for subordinating speech. That was a first step towards my current goal of describing more distinctions to be made between the entitlements that enable subordinating speakers. In my eyes, we shouldn't conflate the ad hoc informal situational authority of the man ranting on the subway with the quasi-epistemic authority of the quack-academic who calmly promotes race-science. Similarly, we should recognize the entitlements afforded by specific relationships or social roles, like being a parent or partner, may also give rise to distinct subordinating speech acts available to different speakers in different contexts. That is, a difference in speaker entitlements explains why the utterance "the thought of you kissing another woman disgusts me"¹²⁵ has a different subordinating force when spoken from a mother to her queer daughter than if that same daughter overheard Pat Robertson saying same-sex kissing "makes [him] want to throw up" (Tashman 2013), though the content is very similar. In each case, the specific position the speaker occupies, including the distinct normative statuses these presume, enable subordinating speech acts with notably different forces.

Acknowledging this type of variability is trivially true in the discussion of slurs. And while not flatly denied by any authors I'm aware of, it remains largely under-examined in the broader subordinating speech literature. I believe it is important to recognize these sorts of distinctions explicitly, and moreover, doing so in a way that does not endorse a rigid hierarchy among these acts, as this betrays the experience of those target by this speech (Gelber and McNamara 2016). After laying out some of these distinctions, I'll argue that McGowan fails at

¹²⁵ Adapted from Ramsey (2013).

the former task (recognizing the distinctions), and Langton at the latter (rejecting a hierarchy of subordinating speech).

2. VARIETY WITHIN SUBORDINATING SPEECH

It's useful to take stock of the different sorts of utterances we've encountered so far, along with some new ones. With these in view, we can note a few rough, intuitive distinctions that demonstrate the different sorts of acts these represent. Doing so will help set the stage for the more in-depth discussion of the various entitlements and outputs that follows.

- a) A South African legislator, in the appropriate circumstances, says: "Blacks are no longer permitted to vote," and therefore makes it the case that black South Africans can no longer vote (Langton 1993, 302–03).
- b) An older white man walks up to an Arab woman on a subway car crowded with people, and says, "F***in' terrorist, go home. We don't need your kind here." He continues speaking in this manner to the woman, who doesn't respond. He speaks loudly enough that everyone in the car hears his words clearly. All other conversations cease. Many of the passengers turn to look at the speaker, but no one interferes. (adapted from Maitra 2012, 100–01)
- c) Outside a store in New Jersey are signs posted reading: "Muslims stay out!" and "We don't serve Muslims!" (adapted from Waldron 2012).
- d) A Black man steps onto a bus filled with only white people, and a white man utters: "Turn around boy, you're not welcome here" (McGowan 2012).
- e) A radio program in Rwanda broadcasts the message "All Tutsi are snakes and must be eliminated"
- f) A man yells from across the street, "I'm gonna kill you f——ts!"
- g) In the comments section on a news website for an article reporting on the shooting death of a young Indigenous man, someone posts: "All Indians are savages!!!"
- h) Late at night, a burning cross is hastily set up outside the home of the lone Black family who recently moved into the neighborhood. The commotion wakes them up. Terrified, they peek outside and see a car pulling away.
- i) Members of the Westboro Baptist Church protest outside the funeral of a soldier who was killed in Iraq, holding signs that say, "Thank God for Dead Soldiers," "God Hates F-gs," and "God Hates the USA/Thank God for 9/11."
- j) "... laziness is a trait in blacks. It really is. I believe that. It's not anything they can control" (Donald Trump, as reported in Cassidy 2018)
- k) A man in the crowd at a football match yells, "Go back to Africa!" at one of the opposing players, who is Black.

- l) A 'White Power' flyer is posted on a number of bulletin boards across a college campus.
- m) A passing car yells "White Power" to a group of young Black children playing outside
- n) A large group of (mostly) white men, marching in a street of US city, chant; "Jews will not replace us!"
- o) Steve and John are co-workers at a factory with very few female employees. They are eating lunch in the lounge and John asks, "How'd it go last night?" Steve responds, "I banged the bitch." John, smiling, asks, "She got a sister?" (Adapted from McGowan 2009, 399)
- p) Susannah says to Jolene, her daughter, "The idea of your kissing another woman is unimaginable and disgusting" (adapted from Ramsey 2013).

As this list of examples helps reminds us, subordinating speech covers many different instances of hateful or otherwise bigoted speech. It can be delivered orally, in writing, on the internet, or even in tangible, highly symbolic form, as the burning cross example shows. It can be primarily ingroup speech, either as an explicit effort to strengthen animosity towards outgroups, or as casually bigoted banter. It can also be speech that crosses identity-lines, either intentionally or not. It can come from a speaker who is known to its target, whether as a personal relation, or as a public figure. And it can, as Delgado and Stefancic say, be "direct and indirect, veiled or overt, single or repeated, backed by authority and power or not, and accompanied by threat of violence or not" (2013, 11). For these reasons I believe it is a fundamentally diverse category, encasing many different types of speech acts.

2.1. *Rough Distinctions*

Before going into detail about differences in the entitlements and outputs, it's useful to first note a few intuitive, though rough, distinctions available. First, there is the distinction, already covered earlier, between assaultive hate speech and propagandistic hate speech. *Assaultive hate speech*, occurs where the speaker directly accosts their target(s); often face-to-face, but can also occur via phone, text, social media, etc. *Propagandistic hate speech*, in

contrast, is standardly ingroup speech about an outgroup. But, of course, outgroup members may indeed be part of the audience of this speech, so we should add: *boundary-crossing propagandistic hate speech*, which is ingroup hate speech that reaches outgroup members, whether directly at the time of utterance, or through other means, like general circulation in the media.¹²⁶

And in all these cases, a relevant difference lies in the fact that the speaker may either be known or unknown to their audience. Public figures are one kind of case, personal acquaintances are another. We can also distinguish between identifiable laypeople who are unknown to the target in one sense, and fully anonymous speakers who are unknown in a different sense. A subordinating speech act uttered by an *Identifiable public figure* would include the South African legislator case, as well as the Donald Trump example above. We could also include the Westboro Baptist Church case here as well, as they are a well-known group, while not holding any official office like the legislator or Trump (eventually). One uttered by a *Personal acquaintance* captures the example of Susanna and her daughter, as well as the factory-lunch example. We can think of *Identifiable layperson* cases as capturing the Subway example, as well as the case where a Black man is accosted while stepping onto the bus, and others. In these cases, the speaker does not occupy a relevant office and isn't known to the target in any other sense either; these are the 'ordinary' instances of hate speech discussed at length in Chapter Two. But *Anonymous speakers* are relevantly different: Where "White Power" or "Muslims Out!" is plastered on a wall, there is no obvious speaker to whom we can attribute the speech act. And

¹²⁶ Gelber and McNamara (2016) refer to these as "generally circulated" instances of hate speech. That is, hate speech events that are reported—whether in the media or by word-of-mouth—to those outside of the original audience.

note how this can occur in direct assaultive hate speech as well, as online harassment demonstrates.

Similarly, the speaker may be alone or part of a group, and this makes a difference for the type of subordinating act being performed. The subordinating effect of one person shouting on a corner seems intuitively than a large *group* chanting: “Jews will not replace us!” And, interestingly, it is within a large group that singularly identifiable individuals may achieve something close to anonymity.

Finally, there is an intuitively relevant distinction is how these speech acts may be singular events, or part of repeating pattern. As Regina Rini argues in the case of microaggressions, a “relatively minor insulting event [is] made disproportionately harmful by taking part in an oppressive pattern of similar insults” (2018a, 332) As she says: “What makes microaggression distinctively harmful is victims’ awareness that each instance is not an isolated accident” (2018a, 335). What’s the case for microaggressions also seems to carry over to other types of subordinating speech, and I explore this more fully below.

This quick roundup of some intuitive distinctions makes a *prima facie* case that there is, in fact, a multitude of speech acts with oppressive force(s), with relevant differences in terms of both the speaker’s role in shaping those forces, as well as how these utterances impact their targets. Going forward, my primary goal is to clarify some of these distinctions using the concepts of entitlement conditions and outputs of normative status. As will become clear by the end of the next section, I understand speaker authority as a subset of broader entitlements, and authority itself can be helpfully distinguished in a few ways.

3. A (ROUGH) TAXONOMY OF SUBORDINATING SPEECH ACTS

As I said at the outset, a complete and precise taxonomy of subordinating speech is an elusive goal. But making some heuristic distinctions about speaker-entitlement can clarify what enables certain speech acts to have the subordinating force they do. Similarly, understanding the different outputs different speech acts typically strive for can reveal the interlocking way that subordination functions. The analysis of oppressive speech is improved when we keep both features in view. So, it is worth tackling this project, even if it will remain incomplete.

3.1. Authority

Authority has played such a large role in shaping the philosophical literature on subordinating speech, so it's useful to dig into this concept first. To be useful however, authority must to be distinguished in more ways beyond just formal and informal. At the general level, I agree with Tirrell when she says, “[d]iscursive authority is a situational power to make felicitous speech acts and gain a range of appropriate uptakes. A speaker’s discursive authority renders her speech acts socially meaningful” (2018b, 15). And, as experience tells us, this situational power can be exercised in many ways.

A conception of formal authority can easily explain why the current President has some significant powers, and why their (executive) orders have the force to compel action that they do. But even here the importance of recognizing *domains of authority* is necessary, as this explains why there are many areas outside of the domain of this authority.¹²⁷ This is explained by

¹²⁷ Nearly all authority is limited to a domain, and so specifying its domain is a crucial task. The one exception is perhaps God’s authority in some religion. As Tirrell notes “Theists ascribe *perfect* authority to God, who, as the original speaker of performative utterances, has creative as well as absolute coercive power” (2018b, 16).

appealing to the formal process that grounds this authority. In the case of the President, there is an inauguration and the constitution, which explains both how one attains these powers, and what their limits are. This type of authority is properly recognized as *institutional authority*, and it is by inhabiting institutional roles that some speakers are authorized to perform certain kinds of speech acts (Tirrell 2018b, 17). Once we identify what role within a given institution a speaker occupies, we learn a great deal about their ability to make authoritative utterances. And, of course, this kind of authority can be used to subordinate. Therefore we should take note of:

- *Formal positional authority*: The explicit, institutionally backed authority held by identifiable speakers in formal positions that grant them power to issue binding directives in a given domain.

This is the type of authority demonstrated in examples such as the South African legislator, and others where the legal rights and permissions are at issue. The domain and jurisdiction of formal authority is, generally, explicitly codified.

But when we're considering informal, or non-institutional, conceptions of authority, things are much less clear cut. The mechanisms by which one gains authority, along with the specific contours of the scope of its domain or the strength of its edicts are much less obvious.

One promising direction is to understand authority as something that an audience can bestow on a speaker.¹²⁸ As we saw in the last chapter, speakers can make utterances whereby they attempt to move themselves into a position of authority over their audience, and how an audience responds can play a significant role in whether or not this move succeeds. So, one option for gaining informal authority may occur via what Ishani Maitra (2012) calls "licensing."

¹²⁸ Maitra (2012) offers one account of how this can occur. Langton's newer work (2017; 2018a; 2018b) also develops this and combines the notion of informal authority along with that of accommodation. Tirrell (2018b) has helpful insights on this topic, as do Ayala and Vasilyeva (2014).

While I raised some complications with her account of this process, I nonetheless agree with the general assessment that a speaker who seemingly lacks authority prior to speaking may gain it when their speech acts go unchallenged. Though, as I argued earlier, background norms play a significant role in determining who can attain this authority, and what the appropriate domains of that authority include. This gives us:

- *Informal situational authority*: The authority that can arise when speakers stake out an ad hoc authoritative position in a given situation, and the audience grants the speaker authority over a specific domain (e.g., appropriate behavior in the local environment).

This type of authority is situational, and as such is not stable across contexts. So, unlike the legislator who usually retains authority over their domain from one day to the next, speakers with situational authority might lose it soon after they acquired it.

But, once we recognize the importance of domain in anchoring authority, we can notice a distinct type of *informal* authority that is nonetheless fairly stable. For a first pass at this idea, consider how something like authority seems plausible when there is a distinct lack of rivals in a particular domain. Utterances that occur within a relatively closed community, like online ‘incel’ forums, can seem to take place apart from the rest of the world. It is in these cases, Langton suggests, “a speaker may have authority relative to one field, where there are no rivals, but [...] the speaker’s authority may be invisible to those looking from outside the bubble” (2017, 18). That is, in some contexts, a vacuum of authority may be exploited by otherwise non-authoritative speakers. Here, a “low type may be all one has, when it comes to authority (Langton 2018a). We may then note:

- *Informal stable authority*: The authority a speaker can have in a specific domain despite lacking institutional sanction, arising from a complete lack of rivals.

Where there is a complete lack of sexual education in schools, pornography may take on authoritative status in the domain of sex, if only for a subset of its whole audience, who then form their beliefs and take their cues from this source. Absent rival candidates for authority, these speakers' words have the performative force of authoritative judgments.

However, appeals to domain only go so far. It may make some sense to claim that pornographers have *informal authority* in the domain of sex because of its near-monopoly status for many adolescents (Langton 2018a, 32ff). But it strains meaning to say that bigoted propagandists, like the white supremacist Richard Spencer, have authority in a domain like immigration because of an absence of comparative rivals. Here, there are rivals, and it's simply not the case that his racist views are the only ones his audience has encountered. Indeed, it's often *because* these views push back against what his audience sees as 'the liberal orthodoxy of the elites' that he—or speakers like him—are given uptake. And so, this explanation alone will not do, and we'll need to add to it. Yet, since a not-insignificant portion of their audience treats them as authoritative in this way, there is clearly something to be said here.

Note, though, how in both cases of informal authority speakers rely on their audience to *recognize* their authority and *respond* with the appropriate uptake to achieve this status. Thus, in informal cases we must note how “[a]uthority results in uptake, and uptake further entrenches authority” (Tirrell 2018b, 17). There is therefore a reciprocal relationship between uptake and authority. To get clearer on this, it's necessary to examine in more detail exactly what we mean when we're inquiring about authority for subordinating speech. This entails clarifying authority in a few key ways, namely, how it operates in a nonideal fashion, and how it may be relative along a number of axes.

3.1.1. *Nonideal authority*

It's useful to distinguish between practical and epistemic authority, as this helps to reveal the nonideal nature of the authority that concerns us. Practical authority is authority for action, while epistemic authority is authority for belief.¹²⁹ The characteristic speech act of practical authority is the Austinean 'exercitive,' which is "a directive, or an enactment of status" (Langton 2018a). What an exercitive does is alter facts about what is permissible and impermissible in a given contexts. It makes some future actions likely, and some improbable. "The law," Langton says, "is a paradigm practical authority, and commands are its paradigm speech acts" (2018a). The characteristic speech act of epistemic authority is the 'verdictive,' which is an authoritative judgment that something is the case. Judges rendering their verdict is a paradigm example, as is an umpire calling a strike.

While analytically separable, practical and epistemic authority are usually exercised together. When the judge gives her verdict, for example, this both declares that something is the case—the fact that the defendant is innocent—and also makes certain actions permissible—that she may now walk free.¹³⁰ But the distinction between practical and epistemic authority helps to reveal the nonideal nature of the type of authority that concerns our topic. The conception of epistemic authority we're interested in is more about *credibility* than it is about actual *expertise*.¹³¹ While one does not become an expert simply because some people take them to be one, being taken as credible is often enough to get others to believe and act on those beliefs.

¹²⁹ Tirrell, in a similar though distinct way, distinguishes between "positional authority," which "grants the holder specific game-assigned powers" relative to "domain in which positions are allocated," and "expertise authority," which "derives from knowledge or skill" (2018b, 17).

¹³⁰ As Langton puts it: "doctor's orders have their status as directives, in part because of what your doctor knows, or is taken to know: his practical authority has its source in his epistemic, or 'theoretical', authority" (2018a).

¹³¹ This is therefore different that the account of "expert authority" developed in Tirrell (2018b).

Large swaths of the wellness industry would not exist were it not for the credibility—not expertise—of those who promote a variety of cures with little to no actual evidence. Similarly, for practical authority, we may recognize that “being put ‘in charge’ may be enough for being in charge” (Langton 2018a). This distinction between credibility and expertise helps to clarify that *de facto* authority is often enough to compel action, at least in this nonideal world.¹³²

This approach makes sense since our topic is not authority as it ought to exist, but authority as it works in the real world. This means moving away from the restrictive conception of formal speaker authority as it is depicted in Austin where, as Langton notes, “authority is often a formal matter: it is the authorized official who can name the ship, conduct the wedding, issue the sentence.” In contrast to this, Langton correctly notes “the workings of informal authority are more subtle” (2017, 33–34).

As we’ve already noted, authority is most often relative to a certain domain. This simply means that there are some subject areas that fall within this speaker’s scope of authority, and some that fall outside of it. Authority being relative to a certain *jurisdiction* amounts to noting how it may be limited to certain audiences. So, some, but maybe not all who hear its call will properly feel its pull. This is what a more on-the-ground, nonideal conception of authority permits. It is, in a sense, a descriptive account of who takes some speakers to be authoritative, rather than a prescriptive account of who ought to be authoritative. So, while the law may be the official authority with the widest scope in a nation, within smaller communities there exists pockets of leaders with greater authority than the law among their enthusiasts. What the Pope or

¹³² The conception authority at issue in cases of subordinating speech is therefore not the same as that which concerns political philosophers interested in the legitimate grounds of authority, that is, that which could morally justify the exercise of authority (particularly of a state), rather than mere *de facto* authority. As Langton says, and I concur: “Our topic is not idealized authority, but something closer to home: structures of social authority, relative to practices, which enable the enactment of norms and hierarchies that are socially real—whether or not they exist in Plato’s heaven, or the better neighborhoods of Earth” (2018a).

a local pastor says may have no grip on the atheist, but it can amount to a clear directive with clear authority to the devoted.

As Langton says, appealing to the idea of a relativity to jurisdiction is helpful because it:

explains how a low type relative to one hierarchy could be, so to speak, a high type relative to another. That might be because being higher in the one system *makes* you lower in the other, and vice versa. One would expect this to hold especially in contexts of conflict, and polarized hatred. Austin's low type, who tries to name the ship after Stalin, may be a leader among his fellow revolutionaries. The 'marginal' figures despised by the majority may, for that reason, be esteemed figures among the minority. The higher one's credibility in the fraternity of doctors, the lower one's credibility in the fraternity of quack doctors, and *vice versa*. The lower one's standing in the eyes of The Great Satan, the higher one's standing on the side of the angels. (2018a)

For this reason, appeals to the 'fringe' status of a speaker might not say much about whether that speaker is an authority in their community, as there is sometimes a sort of inverse relationship between competing communities.

The type of authority at issue is more about how speech is received by its hearers—its uptake—than by objective facts about the speaker's formal social position. It's a type of authority that recognizes that an anti-vaxxer blogger can be authority for millions of parents, even though they lack all traditional markers of authority and expertise on the subject. The fact that it functions like traditional authority for a portion of its audience is the chief reason to still label this capacity *authority*. In other words, if some people take the "quack" doctor's advice as credible, and shape their actions in accordance with their directives, we ought to understand their actions as responses to authoritative utterances, just like the person who takes the "legitimate" doctor's words as sufficient reason for belief and action. In both cases, someone "grants authority to experts' speech acts by taking up the licenses they issue, while deferring responsibility for justification back to them" (Tirrell 2018b, 19).

And what explains the authority of ‘quack’ doctors may also explain some speakers of subordinating speech. The authority the bigot sees in the white nationalist’s words may be as ineffective to some as the Pope’s words are to the atheist, and just as invisible. And yet, within the community of bigots, that speaker might have something close to absolute authority over these subjects. Combining our notions of nonideal authority being relative to a domain and jurisdiction, we’re led to a new conception of *informal stable authority*:

- *Informal stable authority*: The authority a speaker can have in a specific domain, relative to a certain audience, despite lacking institutional sanction. This may arise from a complete lack of rivals, but also where there is competition among competing authorities, and the actions of some reveal that they defer to one more than another.

This type of authoritative entitlement can be found in various forms of subordinating speech. It functions most clearly in propagandistic hate speech as it functions as ingroup speech. However, it may also apply to what I called earlier *boundary-crossing propagandistic hate speech*. Indeed, this speech is often harmful because marginalized people understand that there is a group to whom this speech is like the Gospel. And as the discussion in the previous chapter showed, authoritative speech is about shaping the normative landscape for one’s audience (and other third parties) and this can occur without total audience agreement. Rather than agreement, what it requires is an adjustment that takes on board what speakers presuppose. And while it is often only *part* of their community that does so, the factions that do say a lot about what hierarchies in the community are dominant, and thus, what kind of speech act the speaker performed—whether backed by authority or not. As Langton notes, “[t]he social world can sometimes be accommodating, *especially when helped along by background expectations*” (2018a, emphasis

added).¹³³ Given background conditions of racism, sexism, and much more, the social world is often far too accommodating to subordinating speech, as many in its audience acquiesce to its authority. As the large and passionate following of many hateful speakers—like those of Richard Spencer, Alex Jones, Ben Shapiro, QAnon, etc.—show, there is often a significant portion of people who play the part of a willing, and submissive audience. By responding in this way—by giving this uptake—an audience can demonstrate how a hate speaker’s words are backed up by authority, as this authority is embodied in their uptake.¹³⁴

As a result, an updated, non-exhaustive list of the entitlements for subordinating speech acts would include:

- *Formal authority*: The explicit, institutionally backed authority held by identifiable speakers in formal positions that grant them power to issue binding directives in a given domain.
- *Informal stable authority*: The authority a speaker can have in a specific domain, relative to a certain audience, despite lacking institutional sanction. This may arise from a complete lack of rivals, but also where there is competition among competing authorities, and the actions of some reveal that they defer to one more than another.
- *Informal situational authority*: The authority that can arise when speakers stake out an ad hoc authoritative position in a given situation, and the audience grants the speaker authority over a specific domain.

¹³³ As Langton clarifies: “Accommodation is only a tendency, as Lewis said. It does not always work, and whether it works is itself sensitive to background social hierarchies” (2018a).

¹³⁴ As Tirrell (2018b) puts it: “Authority is largely constituted by audience uptake, which is a matter of next moves (21); and she refers to Bourdieu (1991, 116) who explains: “The symbolic efficacy of words is exercised only in so far as the person subjected to it recognizes the person who exercises it as authorized to do so, or, what amounts to the same thing, only in so far as he fails to realize that, in submitting to it, he himself has contributed, through his recognition, to its establishment.”

3.1.2. *Social (not local) accommodation: Collective speech acts*

As the preceding demonstrates, I believe the uptake given by the hearers of a subordinating speech act plays a significant role in constituting the authority of a speech act. In this, I am not alone. Langton considers audience accommodation to be the main mechanism for how informal authority is accrued by speakers of hate speech. “Authority,” she claims, “can be obtained by *accommodation*, a default adjustment that occurs, without fuss, when hearers take on board what speakers presuppose” (2018a).¹³⁵ In this manner, accommodation “has the power to alter the illocutionary force of an utterance” and, “hearers routinely accommodate what speakers presuppose, and the hearer’s omission is a quiet engine of the speaker’s success” (2018a; 2017, 4).

However, I do not believe the specific mechanisms of accommodation that Langton describes are up to the task of explaining the authority of hate speech—or, at least not the bulk of it. In an effort to clarify the hearer’s role in helping the speaker, I will now point out the limits of her conception of accommodation. As I see it, it is much more active than Langton suggests, and much more social.

The core idea of Langton’s conception of accommodation is that when a speaker says something that requires an adjustment in the shared presuppositions that guide the conversation, this occurs by default.¹³⁶ For example, this sort of accommodation occurs when a hearer reacts to a speaker saying, “Even George could win,” by taking on board the presupposition that George is

¹³⁵ This builds on work by Lewis (1979), who argued that conversations follow a “rule of accommodation.”

¹³⁶ With some qualifications, these shared presuppositions form the “score” of the conversation, which can be thought of as “an abstract structure comparable to the score of a game, which tracks what has been done at a given point, and what it is legitimate to do thereafter” (Langton 2018a).

an unpromising candidate, and future moves within this conversation are explained by the fact that hearers passively accept this low ranking of George.

She suggests, then, that this model may be straightforwardly applied to authority as well, including the authority of hate speech. “Speech acts,” she says, “including directives generally, and hate speech specifically, can acquire authority by an everyday piece of social magic: authority gets presupposed, and hearers let it go through, following a rule of accommodation.” (Langton, 2018b). But extending this idea to the accommodation of *authority* is more difficult than Langton suggests if we want to capture a wider set of cases.¹³⁷

In short, Langton’s model moves too quickly. What’s required is a more detailed understanding of the multiple audiences a speech act can have, along with greater appreciation for how an audience can itself *actively* aid in the construction of authority, and not simply passively accept it. Once we acknowledge these features and modify the conception of accommodation to properly capture what is occurring in these cases, we’re led to a novel type of speaker authority that I call *collectively sustained authority*.

Consider Richard Spencer again, who rose to prominence by spearheading an unabashed white supremacist ideology. At the height of his popularity, he had legions of followers, some of whom he led in a Sieg Heil salute while he *and the crowd* chanted, “Hail Trump!” (Wood 2017).

What’s at issue here, in my eyes, is not accommodation in the conversational sense as described by Lewis (1979) and taken up by Langton, but rather accommodation in a much more extended social sense. The type of accommodation that a willing audience gives authority to hate

¹³⁷ Langton does acknowledge differences between these two types of accommodation, but for her it is mainly a difference in effect (acceptability vs. truth), and the mechanism is largely the same: “Unblocked presuppositions tend to become acceptable rather than true: ‘George is an unpromising candidate’ may become acceptable, if unblocked, but that would not make it true. Authority is different. A presupposition of authority can become, not just acceptable, but true, because its existence, not only its acceptability, depends in part on what hearers do, or fail to do.” (Langton 2018b).

speakers is, I claim, significantly different than the accommodation that occurs within a single, relatively shorter conversation. It therefore demands a different explanation, specifically one that allows for the active participation of a willing audience to bolster its claims to authority, even in the presence of sustained attempts by outsiders (e.g., anti-racist activists) to quash these group through speech and other means. In other words, we must recognize that counter-speech doesn't do the job of preventing authoritative speech in the way the conversation model describes.¹³⁸

I do agree with Langton, however, that it is largely because of the audience-participation that this type of subordinating speech can be called authoritative and has the specific subordinating force that it does. But rather than focus on the segment of the audience who let these utterances pass by without much notice—who either take on board their noxious presuppositions or at least fail to block them—I want to emphasize the role of those who take up these utterances willingly, and actively endorse them as authoritative, and in doing so expand the reach of these speakers even wider with their own speech acts.

First, it's worth acknowledging that the rise to authoritative status of figures like these often takes time.¹³⁹ And even when it is or feels sudden, it's still much longer than a single conversation, and so is different than how the hike-planner or the restaurant-chooser attains authority in common examples of accommodation (Maitra 2012; Thomason 1990, quoted in Langton 2017). In this way it's different than the *situational* authority of the hate speaker on the streets, and this makes sense, as the authoritative status of these figures is more stable.

¹³⁸ I take up this last idea in Chapter Four in more detail.

¹³⁹ One illustrative example is found in Jamie Bartlett's *The Dark Net*, where the sudden rise of white nationalist groups in the UK is described to take place over the course of months and years, moving from the online to the streets and back again. See Bartlett (2014, esp. Ch. 2). For other accounts of real life events of the rise (and fall) of prominent speakers in hate movements, see (Strum 1999); (Hategan 2004); and Saslow (2018).

To capture this sort of accommodation we're better off appealing to an idea like the shifting boundaries of the permissible as described by philosophers like Jennifer Saul (2017) and Lynne Tirrell (2012; 2017).¹⁴⁰ Saul picks up on how accommodation enables a hearer to shift their understanding of what's acceptable, for example, how "openly racist utterances effect significant changes to standards of conversational acceptability" (101). Saul draws on accommodation as it is described by Langton and McGowan, but she makes the crucial addition that this shift in permissibility standards can be effectively forced. Through the use of a "diachronic figleaf," a speaker's earlier explicitly racist utterance can be reconfigured as seemingly unproblematic (ibid. 105).¹⁴¹ In explaining how this works, Saul discusses how a speaker and their allies can work to undermine standard inferences about the problematic interpretations of an utterance. And, in doing so, they *force* the accommodation of new norms on an otherwise resistant audience. Furthermore, she notes how "what becomes permissible within one community will not be permissible in another" (108). In doing so, she pushes back against the notion that there is a single, shared score that all participants track.

This sort of audience-contribution is more active than accommodation, and shows how the hearers and—crucially, in my eyes, *repeaters*¹⁴²—of hateful utterances can play an important role in shaping the (discursive) norms of the wider community. They do so actively, not

¹⁴⁰ As Lynne Tirrell says: "Our speech acts also undertake a meta-level expressive commitment about the very saying of what is said. Expressive commitments are commitments to the viability and value of particular ways of talking, modes of discourse" (Tirrell 2017, 144). When these expressive commitments aren't uniformly condemned, they can shift the boundaries of what counts as acceptable discourse in a community, opening the door to further harm (Saul 2017).

¹⁴¹ As Saul explains, a "racial figleaf is an utterance made in addition to one that would otherwise be seen as racist [that] provides cover for what would otherwise have too much potential to be labeled as racist" (Saul 2017, 103). These are, in her view, tools that enable that provide cover to obviously racist statements, rendering them acceptable.

¹⁴² Tirrell discusses a similar type of authority, briefly, when she mentions "parroting," which she says, "has only *deferred* authority, derived from the expertise of the person who ultimately bears justificatory responsibility." She also says that when "a speaker's utterance piggybacks on the authority of another's utterance but moves beyond parroting to extend the claim, this mixes position and expertise, and blends deferred authority with independent authority" (Tirrell 2018b, 19).

passively, with action rather than omission, and with intent, rather than unconsciously. In this way, this discursively forced accommodation is no “quiet engine of the speaker’s success,” but instead an often-loud assist.

This, then, allows us to see how speech acts in addition to those from the original source are also backed up by a type of authority, such as those who chant along to a racist mantra, or who later repost similar messages on internet comment boards. This is a different power not explained by accommodation, as Langton understands it, as her account seems to focus on establishing single speakers as authoritative, rather than the mass of followers who often join in, with acts of repetition.

Because of the reliance this type of subordinating speech on this part of its audience, whom I see as collaborators and promoters, I view these as *collective speech acts*, which rely on a type of *collectively sustained authority*. Its chief feature is the endorsement of a critical mass of supporters who do not so much passively accept the authority of the other speaker(s) as enthusiastically support it. In doing so, they take the speech beyond its environment, and flex its strength over outsiders.

- *Collectively sustained authority*: Through repetition and endorsement, individually non-authoritative speech is made authoritative by signaling collective support and solidarity. A group of speakers have their authority secured through a repetitive process where each new utterance adds to the strength of the overall practice.

This type of subordinating speech has not been fully addressed in the literature, and I examine it in more detail in the following chapters.

Seeing these distinct types of subordinating speaker authority broadens our understanding of the actions and practices that sustain these speech acts. In marshalling these different types of

authority, speakers produce unique subordinating speech acts. This analysis allows us to see both what distinguishes these acts from one another, and what unifies them.

Lastly, it is worth acknowledging how subordinating speech is possible absent any speaker authority. This runs afoul of the strict interpretation of subordinating speech acts require authority because it engages in illocutionary acts like ranking, legitimating, and prohibiting. However, once we broaden our conception of what it means for speech to subordinate and understand it more as speech acts that are part of broader practices of subordination, a wider range of utterances fall into view.

- *Non-authoritative entitlements*: The sort of entitlements arising from *direct relationships* that enable a variety of types of subordinating speech. A direct relationship can be any relationship that allows a speaker to interact with the audience directly, which opens the door to the expression of subordinating speech.

To be clear, these speech acts are not absent specific *entitlements*, but only free of what I'm understanding as authority. For example, a classmate might mock another student for their 'foreign' name or disability; a sibling might express disgust at their gay brother; a random person on the street who stopped you to ask for directions might suddenly ask "where you're *really* from." In cases like these, the speech act has plausible oppressive force less from the speaker exercising any speaker-authority over their addressee, but in how these speakers bring the authority of oppressive ideologies to bear on their targets.

Beyond adding to our collection of subordinating speech acts, recognizing the possibility of non-authoritative subordinating speech serves as an important reminder of the crucial role that oppressive norms and practices play in *all* instances of subordinating speech. In other words, background structures of oppression play an essential but under-recognized role in filling out the story for all these entitlements. It is not always the entitlement aspect of the speech act that

renders it most oppressive. This is only one element. I believe it is an important element, worthy of the sustained investigation I did above. But it is worth acknowledging that in some cases, it is the content that is most potent, considered apart from the speaker, and this much must be acknowledged.

3.2. *Output*

I began by canvassing speaker authority for subordinating speech because this feature plays a significant role in shaping the output of these speech acts. As Tirrell points out, the “success of a speech act relies in part on the authority of the speaker to drive audience uptake. Speaker authority and audience uptake go hand in hand, and neither occurs in a vacuum” (2018b, 13).¹⁴³ We saw this in how the uptake of some either clarified or actively shaped the type of speaker authority in play.

Turning now to the output side of subordinating speech acts, I will examine this output in terms of three features: to whom is it primarily addressed; whom does it target for subordination; and what normative statuses it assigns, which is informed in part by the first two features, along with the specific entitlements it marshals.

3.2.1. *Addressee vs. target*

It is helpful first to distinguish between subordinating speech acts in terms of where they are first *directed*, both in terms of whom they are addressed to and whom they target. Even

¹⁴³ Tirrell notes how she takes “uptake to be a matter of next discursive moves, moves either depending on or disregarding the original speech act” (Tirrell 2018b, 15).

though oppressive utterances ultimately *target* a marginalized group, some are addressed primarily at an individual within that group (e.g., a Black man on the street), others at a small subset of the larger group (e.g., a Black sorority), and still others at an entire group (e.g., all non-white people). While an epithet directed at a single individual will still have reverberating oppressive effects on other members of that group, the way it achieves this effect is different than a hateful comment directed at all members of the group at-large. By noting to whom an utterance is directed, we're in a better position to understand who the speech act calls upon for a response, and thus puts under distinct normative pressure. As Kukla and Lance argue, speech acts are essentially *vocatives*, insisting on recognition and response from their audience. Locating the addressee of subordinating speech is a first clue to identifying its pragmatic output.¹⁴⁴

Furthermore, it is useful here to return to the rough distinction between assaultive hate speech and propagandistic hate speech discussed above. Propagandistic hate speech is typically spoken by members of one group to fellow ingroup members (e.g., a white person to other white people). It is thus, most often, intra-group speech. Assaultive speech, on the other hand, is typically inter-group speech (e.g., a white man accosting an Indigenous woman). Assaultive speech can therefore address an individual in any of the three manners considered above, while intra-group propagandistic speech generally does not address the outgroup members it disparages at all.¹⁴⁵

¹⁴⁴ In particular, they write: “Concrete normative relations among people are established and sustained through vocatives—that is through the Yo-claims that *hold us in place* in social space. We become and remain the types of beings that have specific, agent-relative engagements with others through an ongoing network of hails and acknowledgements. [...] Vocative discourse plays a crucial role in constituting individuals as particular, normatively positioned persons” (2009, 181).

¹⁴⁵ I say generally, in part because issues of intersectionality inevitably arise in certain instances. For one example, consider that while a white supremacist speech will often reserve most of its vitriol for the non-white races, there is likely also much misogynistic content to be found as well.

But even when speech is intra-group, when it is subordinating, it also has the function of *targeting* outgroup members in some form, again possibly as individuals, subsets, or whole groups. Often, this targeting is quite explicit. When ‘incels’—the misogynistic community that self-identifies as ‘involuntary celibates’—discuss women in their (usually online) semi-private communities, they use the derogatory and alienating term ‘femoids,’ making it clear whom their words subordinate without actually addressing any women.¹⁴⁶ However, the reference need not be so explicit. The famous “14 words” recited by white supremacists does not refer to any other group besides white children. In such a case, non-whites are targeted for subordination largely by omission. Something similar occurs in some cases of assaultive hate speech too, as when “white power” is yelled from a passing car of white people towards a group of Black children. The content of this utterance makes no mention of its targets, but it is nonetheless specific in its oppressive effects.

We can therefore note a distinction between the audience of a speech act (to whom it is directed at, as its primary listeners/readers) and the target(s) of a subordinating speech act (to whom it functions as a tool of oppression), while recognizing that the latter is not always identified in either as the direct addressee of an utterance or as part of its content. Putting things together, then, there are at least four ways a speaker can direct their speech towards their target (at an individual, at a subset of group, at the whole group, or not at them at all), and at least three ways in which the target of subordination might appear in the speech act (as its addressee, in its content explicitly, or by exclusion).

¹⁴⁶ For an informative overview of the ‘incel’ ideology, see ContraPoints (2018), the stage name of Natalie Wynn, on YouTube: <https://www.youtube.com/watch?v=fD2briZ6fB0>; or Zack Beauchamp (2019).

Again, I consider this to be a heuristic rather than an exhaustive categorization, and some of its boundaries are more rigid than others. It is perhaps best to think of the Individual/Subset/Whole-Group markers as points on a spectrum. But this offers a more precise analysis of how marginalized individuals or communities are targeted by subordinating speech than the assaultive-propagandistic speech binary. It recognizes more fluidly how these categories may overlap and offers an initial answer for why subordinating speech acts might differ in their outputs. But it isn't yet to discuss these outputs themselves; I turn to that now.

3.2.2. *Output as normative status*

Following Kukla and Lance, we can conceive of the output of a speech act as “the normative changes (in the status of the speaker, or of others in the discursive community) that the act strives to produce” (2009, 15). I understand these as the normatively-laden roles a speech act aims to impose, where these roles are constituted by a variety of expectations, obligations, and permissions of different strengths. When we're considering specifically *subordinating speech* however, the roles distributed are hierarchical, with some given positions of dominance that others are denied. In other words, the normative changes produced via subordinating speech include the distribution of particular roles with unequal relations of authority, esteem, and status.¹⁴⁷ This tends to occur via the assignment of ingroup and outgroup normative roles. These are the main outputs of subordinating speech acts, in all its varieties. But because oppression is a complex phenomenon, I maintain the ways that subordinating speech acts achieve this end, along

¹⁴⁷ Relational Egalitarians, and in particular Elizabeth Anderson, have persuasively argued that hierarchical distributions of authority, esteem, and status are the primary inequalities that ought to give us concern. See Anderson (1999) for an influential argument to this effect.

with the particularities of these expectations, permissions, and obligations, are similarly complex.

My current concern is thus the distribution of hierarchical roles that subordinating speech makes possible, and how this occurs for the target, the speaker, and other parties as well. This focus on *normative* effects differs from the discussion of the (broader) range of harmful effects discussed in Chapter One. And while I argued there that an adequate account of the harms of subordinating speech must account for these in an important way—and therefore, not focus solely or even mainly on illocutionary effects—I do believe that this aspect of output can be fruitfully explored. That is, apart from the wide range of harms subordinating speech produces, it also puts its targets under distinct normative pressure, and this may be captured under the idea of output as a normative status.¹⁴⁸

While I aim to bring out a diversity of outputs, I should clarify this claim. On the one hand, of course different subordinating speech acts enact different statuses. As Tirrell says, “variability is the norm,” (2017, 149) and uniformity would be a surprising result. For this reason, it’s important to go beyond conceptions that lack the necessary nuance that this topic demands. However, on the other hand, no two events will be exactly alike, and so some generalizations are necessary. Yet, while necessary, they are fraught and must be approached carefully. This is especially true on what I’m calling the output side, as individual features play a

¹⁴⁸ Here, it is worthwhile to briefly note how I see speaker intentions coming into play in the output of subordinating speech. I believe intentions fit in here, on the output side, as they shape the extent of the derogation, which I take to be a feature of output. But they aren’t that important, as they do not fix output. Rather, intentions are one element that inflects the overall shape of the output of subordinating speech. A villainous intent, paired with a contemptuous tone, can render nearly any speech act subordinating (e.g., “Of course, *you would* say that.”) But a non-oppression, even helpful intention is in no way a reliable guide to establishing the lack of subordinating force of a speech act. Attending to speaker intention does, however, allow us to categorize subordinating effects as either the aim of a speech act (e.g., assaultive hate speech) or as a misfire on the speaker’s part (e.g., news reports that amplify harmful ideologies). And so, they can be helpful for establishing what speech act occurred, and are not irrelevant. But the importance of intentions can easily be overstated, and I largely leave them aside.

large role in shaping how one interprets and responds to any given utterance. Moreover, given that I understand output in an open-ended way—where it is not fixed in the moment of the utterance but instead is constituted by its specific effects on the normative statuses of speaker, target, and others—further factors always have the potential to shape output in unexpected ways.¹⁴⁹ My aim therefore is to make vivid a fairly basic point: subordinating speech is a complex phenomenon, and reductionist analyses are bound to erase important distinctions. When aiming to capture the lived experiences of real people, however, theorists like myself must differ to testimony of those most directly affected.

With this in mind, my goal in the remainder of this section is to note the range of normative effects subordinating speech produces, first as this has been described by those targeted by it. To do so I rely on Katharine Gelber and Luke McNamara’s (2016) “Evidencing the Harms of Hate Speech,” which draws on 101 interviews with members of Indigenous and other minority ethnic communities in Australia about their experiences with hate speech. They give voice to those targeted by subordinating speech, and in doing so, provide a clarifying lens to the philosophical literature, which too often excludes those with personal experience with hate speech.

It’s useful to note the distinctions Gelber and McNamara draw from these interviews. They categorize different “hate speech events” into two broad types—“face-to-face incidents” and “broadly circulated incidents.” This categorization roughly matches the assaultive speech/propagandistic speech distinction discussed above. At the same time, however, they are clear to note how these “two types of hate speech were not experienced as qualitatively different

¹⁴⁹ But this is no threat to the general claim that subordinating speech subordinates, since “not every instance of hate speech need result in all the harms elucidated for the argument to be sustained.” (Gelber and McNamara 2016, 336).

in terms of seriousness or harmfulness,” and thus claim we must give each equal attention (2016, 336). They further divide each type with five further sub-categories (epithets [verbal and symbolic]; exclusion; negative stereotyping; transmission of racism; and threatening, harassing behavior). So, given this categorization, we’re led to ten different sorts of hate speech incidents, as derived from the target’s description of the experience.

Their considered position is that “both [broad] types of hate speech events can incur harm constitutively, consequentially or simultaneously in both ways,” and so the divisions are somewhat artificial (2016, 326).¹⁵⁰ One conclusion I draw from their analysis of these interviews is an appreciation for the equal seriousness and harmfulness of diverse forms of oppressive speech. And with Gelber and McNamara, I maintain that analyzing the distinctions among subordinating speech is useful for the purposes of understanding different hate speech events, rather than ranking them, in part since their boundaries are unquestionably blurry, and overlap is common. That is, the distinctions to be drawn are not—and ought not be—hierarchical.

Even within these categories, important differences are to be expected at the level of output. Consider the normative effects a subordinating speech act might have when the direct addressee is the same as its target of subordination. These are categorized as face-to-face incidents by Gelber and McNamara, and this is the situation critical race theorists Mari Matsuda, Kimberlé Crenshaw, Richard Delgado, and Charles Lawrence III describe as assaultive speech, in their *Words that Wound* (1993). Beyond being assaulted, victims also describe these situations as ones in which they are embarrassed, humiliated, ambushed, terrorized, harassed, threatened,

¹⁵⁰ They explain this in the following: “We argue that the distinction between face-to-face encounters and general circulation hate speech is not always clear in the everyday experiences of racism endured by targets. For example, hate speech yelled at a target from a moving vehicle or on the street constitutes a face-to-face encounter, yet can also be experienced as targeted at the community to which the target is perceived to belong. Even if no-one else heard the hate speech at the time, others will hear of it through word of mouth, community events or the media. They thereby become informed that racism is alive and well, and that they may become its target.” (2016, 326)

scared, and so on. These are, on the standard framework, most often seen as perlocutionary effects. But these terms also imply the designation of a status, constituted by permissions, obligations, and expectations. And these can have varying strengths.

We can see this when we examine incidents of ‘exclusion,’ as described by Gelber and McNamara’s interviewees. Many reported permutations of being told they were not welcome and should “go back to where you belong” (2016, 329). In their words, part of the harm consists in how “you feel ostracized”; another reports the pain of “that fear of being judged, of that fear of being sort of like prosecuted, being excluded from the nation’s society” (2016, 334). As I interpret these statements, part of the harm of these incidents seemingly results from being accused of violating an expectation—one that they could not ever meet. Whether it is because ethnic minorities are not assimilating ‘properly’ or are simply in the country at all, exclusionary hate speech expresses indignation at the perceived transgressions of marginalized people. They are thus second-personal calls, and it is in this way that they “institute and reconfigure normative statuses and relationships” (Kukla and Lance 2013, 457).

Depending on the specific circumstances of the overall discursive context, different instances of direct exclusionary speech will shape a new normative status for its target in varying ways. In some cases, this may be as a literal target for violence, and so they will feel threatened, ambushed, and harassed—simply because of their existence. This can occur when the speaker is angry, in a physically aggressive pose, in a group, or otherwise intimidating their target.¹⁵¹ In other cases, the utterance may be less-than threatening, but nonetheless perform a clear

¹⁵¹ Children are especially vulnerable to this sort of exclusionary speech. For example, Gelber and McNamara relate how one interviewee’s “child at school was told by another child ‘The fucking Indians, you can go back to your country’” (2016, 329). This demonstrates how the power a speaker can possess over the target is relative to the target in important ways—an eight-year old can seriously intimidate other children the same age or younger, but will likely have trouble accosting a fifteen-year old.

ascription of inferior qualities—marking one as an ‘unwelcome’ addition to the existing community. For instance, if a work-colleague makes a fuss about being unable to pronounce their addressee’s “weird” name.¹⁵² In either case, the speaker’s utterance has the force of assigning certain context-specific expectations and obligations of different sorts to the target, and this is achieved via the designation of a status. In the former it is as ‘unwelcome’ in an existential way, and so the only practical solution is to leave the situation. In the latter it is as an ‘unwelcome’ and burdensome change to the status quo, and so there may be an expectation of *deference* to the speaker, to not make ‘any more of a fuss.’ So, even within the status imposed by exclusionary speech we see differentiations in how this may be experienced as a normative status. I maintain that similar points can be made about other forms of subordinating speech, like ‘negative stereotyping’ speech, and more.

My position, therefore, is that while differences at the level of form—to whom is the speech directed; who does it target for subordination—aid the analysis of determining who in the community may have their status impacted as the output of a subordinating speech act, there will nonetheless be idiosyncrasies that make a precise and deterministic analysis of output somewhat hopeless. This follows, I believe, from understanding oppression as a complex and multiply realizable phenomena.

But this is not to say that useful analyses of output are entirely elusive. Just because we cannot give a precise analysis of the normative statuses imposed via subordinating speech is not to say there is nothing to be gained by examining the output of subordinating speech. We see this

¹⁵² And, as is obvious to many already, what counts as ‘weird’ has undoubtable racial elements to it, as one interviewee makes clear: ‘So if a French pronounces in their French accent, that’s okay [...] but when an Indian pronounces a word a bit differently to [...] an Australian [...] then it becomes a problem’” (Gelber and McNamara 2016, 329).

clearly once we recognize that the target is not the only one whose status is affected by the speech act. And here, more precision is perhaps possible. As Popa-Wyatt and Wyatt note, “the speaker imposes not just one, but a complementary pair of discourse roles on the conversational score: the speaker takes on a powerful, dominant role, while assigning the target a subordinate role” (2018, 2894–95).¹⁵³ That is, subordinating speech shapes the normative status of not just the target, but the speaker as well.

We see this in the telling use of the verbs “judged” and “prosecuted” by one of Gelber and McNamara’s interviewees quoted above. It is telling in that it lends support to the view that hate speech has an authoritative dimension.¹⁵⁴ But we cannot assume that this authoritative element is all the same sort. Nevertheless, the assignment of dual roles, with corresponding statuses, is a helpful conception of interpersonal subordinating speech acts and can be applied to other types of speech acts as well.

For example, consider a case where the subordinating targets are not addressed directly. Here too, there is still an imposition of hierarchical roles. However, in this case it is not simply between speaker and addressee, but one where speaker and addressee are on one side, with their subordinating targets on the other—ingroup and outgroup.¹⁵⁵ As Langton says, “propagandistic hate speech aggrandizes the hearer, often, as well as the speaker. Propaganda’s message of *their* ‘inferiority’ [...] is also a message of ‘our’ superiority, the superiority of a ‘we’ that includes

¹⁵³ Popa-Wyatt and Wyatt use the term ‘discourse roles’ to here, though I put that aside mainly to avoid having to explicate more of their framework than I think is necessary. Though I should acknowledge that Popa-Wyatt and Wyatt are speaking specifically about slurs and slurring acts. They claim: “slurs are constitutive of oppressive speech with respect to the discourse roles, because they have the illocutionary force of assigning a subordinate role for the purposes and duration of the discourse” (Popa-Wyatt and Wyatt 2018, 2895). But I believe this is a helpful understanding, applicable to other subordinating speech acts as well.

¹⁵⁴ As Gelber and McNamara note: “the harms attested to by interviewees—as experienced, perceived and feared—bear a close resemblance to the harms alleged in the literature, and are both constitutive and consequential” (2016, 336).

¹⁵⁵ See Herbert and Kukla (2016) for a useful discussion of discursive practices of ingrouping and outgrouping.

both speaker and hearer” (2018a).¹⁵⁶ Anti-immigrant speech offers a clear example of this, where the ‘threat’ embodied by (usually, and sometimes explicitly, non-white) migrants is that they will disrupt and otherwise undermine a stable, prosperous community. In this way, the targets of subordination are cast in an inferior role, while the ingroup addressees are reified alongside the speaker. This is the sometimes-seductive aspect of propaganda that we must keep in view, and that I believe the role-assignment conception of output does well to clarify.

Seeing the output of subordinating speech acts as a distribution of hierarchical roles, we see how speech shapes the normative contours of relations between social groups. Whether directed at outgroup members in intergroup speech, or at ingroup members in intragroup speech, subordinating speech has the pragmatic effect of elevating the speaker’s status, while usually at the same time denigrating the status of those that it either explicitly targets or merely excludes. The distinctions uncovered here—different ways of being addressed; different ways of being targeted—are a useful clue in determining who is captured within the output of a subordinating speech act, and how.

Moreover, recognizing the type of entitlement being put to use shapes the sort of output the speech act generates, and so offers insight into how these may be resisted. While practical and epistemic authority often come together, they each call for a different sort of challenge. That is, since they rely on different grounds they generate distinct effects, and “resistance or support may also involve distinct kinds of challenges” (Tirrell 2018b, 18). Practical authority is open to direct resistance through disobedience, while epistemic authority may require counterevidence to

¹⁵⁶ Popa-Wyatt and Wyatt acknowledge something similar when they note how “an act of assigning a new discourse role can be used to shape the power dynamics in a conversation. Discourse role assignment is thus a meta-move in the conversational game in that it changes the discourse rules. [...] Discourse roles are not solitary. A role assigned to one participant typically presupposes complementary roles, to be adopted by other participants. Thus, by entering into a role, a speaker can invite or compel other participants to take on new roles” (Popa-Wyatt and Wyatt 2018, 2892)

undo. This, perhaps, offers a clue as to when and where ‘more speech’ may be useful, and where it is misguided.

Along with the various entitlements speakers might possess, these distinctions in output lend credence to the claim that multiple types of subordinating speech acts are to be expected, and that a reduction to a single type is too simplistic to capture the reality of the phenomena in question. With that concern in mind, I next make explicit how alternative models offered by Langton and McGowan fail at this task.

4. A HIERARCHY OF SUBORDINATING SPEECH

Despite the apparent variety of subordinating speech acts, dominant approaches in the literature seem to neglect this fact. In place of recognizing a rich diversity, McGowan’s model takes all oppressive speech acts to be exercitives. And while Langton is careful to distinguish between different origins of a speaker’s authority to subordinate, her account nonetheless relies on an implicit and monolithic hierarchy of subordinating speech acts, flattening their differences into merely one of degrees. Both are worth considering, and ultimately rejecting.

This latter approach is evident in Langton’s classic paper “Speech Acts and Unspeakable Acts” (1993), where she argues that the speech of pornographers itself constitutes a subordinating speech act, and she does so by analogizing it to the speech of a legislator in apartheid-era South Africa. Since it is only authoritative speech that can to subordinate in the way Langton discusses, she supports her claim by showing that the speech of pornographers is more like that of a legislator than that of dissident—more like the law than protest. Thus, her

strategy is to *approximate authoritative speech*. I'll consider this in more detail in section 4.1 below.

A different strategy is found in the work of McGowan, who rejects the necessity of authority for subordination (2003; 2004; 2009; 2012). Rather than seeing the potential to subordinate (with speech) as something that resides in individual speakers, McGowan sees the ubiquitous nature of systems of oppression as generating an ever-present normative landscape, where individuals make moves that alter this landscape. On this view, the entitlement conditions required for oppressive speech are rather thin. It is the system itself where the power to subordinate resides, not the speakers. I believe there is something very insightful about this account, so it is worth considering in more detail.

McGowan develops what she calls the *covert exercitive* to explain speech acts that function like Austinean exercitives without being fully explicit, and without requiring speaker authority. This has the benefit of explaining the widespread nature of oppressive speech, but, as I argue, this neglects the differences that exist between speakers and speech acts. In short, this model is appealing for explaining (part of) the oppressive force of utterances like microaggressions and other 'casually' bigoted comments. But it is lacking when it comes to accounting for the oppression-enhancing features embodied by some speakers and not others. I consider this framework for understanding oppressive speech in section 4.2 below.

Despite their difference, I argue that both Langton and McGowan's account are ultimately reductive, and that both see subordinating speech as derivative of other speech acts. Because of this, they therefore fail to properly acknowledge the variety of types of subordinating speech acts and collapse this rich category in unhelpful ways.

4.1. Approximating Authoritative Speech

In “Speech Act and Unspeakable Acts,” Langton says the speech act of the apartheid-era South African legislator offers a “clear example” of what the paradigm for subordinating speech might be (1993, 310).¹⁵⁷ In Chapter One I argued that this example is not, in fact, a particularly clear or helpful example of subordinating speech. This is the case because Langton leaves background conditions that enable this utterance to both cause and constitute subordination unexamined. Here, I want to address a different aspect of Langton’s use of this example that distorts our understanding of the broader field of subordinating speech acts, namely, its reification of a narrow conception of authority.

The reason the apartheid-era legislator’s speech is a “clear example” of subordinating speech, Langton says, is that the legislator clearly occupies a position of authority within the formal South African legal context. It is this position, in this context, that provides the speaker with important powers over their targets.¹⁵⁸ And because the connection between authority, formal social position, and subordinating speech is so apparent here, Langton uses this example to build a model of different forms of subordinating speech, such as pornography (1993; 2017) and—in later work—racist hate speech (2012; 2018a; 2018b). Given this methodology, the claim is that these forms of subordinating speech can only approximate the “clear example,” because their speakers can only approximate the paradigm of speaker authority embodied by a legislator.

¹⁵⁷ As I note in Chapter One, she does not explicitly call the legislators utterance the paradigm. Specifically, she writes: “I have not tried to say exactly what the paradigm for subordination is, but I have suggested that the speech acts of apartheid offer a clear example” (1993, 310).

¹⁵⁸ In Chapter One, I argued that the importance given to speaker-position here is over-emphasized at the expense of other, equally important background factors that are also necessary for the subordinating speech to function.

This idea of *approximating authoritative speech* is apparent in one of the main examples Langton uses in more recent work to analogize to hate speech. This is an “Alabama slave law” that, in 1861, said:

that slaves are people and things: as ‘rational human beings, they are capable of committing crimes; and in reference to acts which are crimes, are regarded as persons,’ but ‘because they are slaves, they are incapable of performing civil acts, and, in reference to all such, they are things, not persons.’ (Langton 2018a, citing Caterall 1926, 247)

As Langton notes, such a law performs some significant acts. It deprives enslaved persons from performing important civil acts, like testifying in court, in part by ranking them as inferior from the white population.¹⁵⁹ Notably, these legislative acts are very much like what Langton describes as the main functions of hate speech. Yet, for Langton, the question “is not whether a law could be hate speech. It is “whether hate speech could be like a law: whether hate speech could have authority, and if so, of what kind, and from what source” (2018a).¹⁶⁰ The implicit idea throughout Langton’s work is that subordinating speech will resemble subordinating *law* in a significant way, while not quite attaining the type of authority embodied in cases of the Alabama slave law, or the South African legislators pronouncement.

For this reason, I dub this approach to subordinating speech the ‘approximating authority’ model. The main feature of this model is a combination of three claims: First, that authority is required for a speaker to be able to perform certain subordinating acts with their speech;¹⁶¹

¹⁵⁹ It’s important not to overstate the power of speech here. As Catherine MacKinnon would say, the law is an instance of “authoritatively saying someone is inferior,” which, along with “material supports,” she says is “how indifference to violence against those on the bottom is rationalized and normalized” (1993, 31).

¹⁶⁰ Langton takes a similar approach in another more recent article that addresses the authority of pornography. The topic of that paper is embodied in its title, “Is Pornography like the Law?” See Langton (2017).

¹⁶¹ I say “certain subordinating acts” because I want to note that it is only for certain, possibly paradigmatic, subordinating speech acts that authors such as Langton say authority is required. She acknowledges that speech may harm in many ways absent authority, noting that “Unlike the law, the harm in hate speech dwells partly in factors independent of its authority.” She adds, however, that “Whatever else ‘the harm in hate speech’, it will be the worse when hate speech has authority,” as “authority would surely exacerbate” whatever problems hate speech presents independent of authority. See Langton (2018a)

second, that subordinating authority is located primarily in individual speakers; and third, that different speakers can approximate authoritative subordinating speech to differing degrees. Together, these claims constitute an account of subordinating speech that emphasizes the centrality of individual speakers of differing degrees of (subordinating) authority. When we're wondering whether a given utterance is a case of subordinating speech, this model leads us to ask whether that speaker occupies an authoritative position in the domain under discussion.

4.1.1. What Langton's account leaves out

The main problem with Langton's (updated) model is that it nonetheless remains monolithic in certain ways and retains a hierarchical, individualist picture of authority. And I believe these features limit its applicability to a wider range of subordinating speech acts, as discussed earlier in this chapter.

I say this approach is still monolithic, despite its internal divisions, because it holds onto the idea that speech acts can be assessed in terms of their strength in a straightforward way. That is, for Langton, the issue comes down to whether an utterance is plausibly backed by authority. That authority may be practical or epistemic, and it may apply only to a particular domain or jurisdiction, and it may gain much of its authority simply by a lack of competitors, but the question remains one of asking whether it makes sense to attribute authority to an utterance or not. This approach retains an overtly hierarchal picture of subordinating authority. While there are a few new dimensions of authority, the overall picture remains the same. Some utterances are clear instances of authoritative speech—e.g., the law, expert testimony—and subordinating speech approaches these paradigms to greater or lesser degrees. In this way, it holds onto the idea

that subordinating speech is *more* harmful the more it approaches these paradigms of authoritative speech. And I'm not convinced this is the correct approach to take for subordinating speech.

First, (as I claimed in Chapter One) we ought to reject the idea that there is a single paradigm of subordination, but this idea seems to be central to Langton's model. Indeed, her touchstone examples—the apartheid-era legislator, the Alabama slave law, Nazi propaganda—are positioned as themselves paradigm instances of subordinating speech. And while they are each horrific examples of the category, I worry that this reinforces an explicit hierarchy of subordinating speech, and therefore positions much of this category as less harmful and less significant. Seeing subordinating speech as a category that varies horizontally and not just vertically, as I propose, would do some work against the knee-jerk dismissal of 'less severe' forms of subordinating speech than outright legal rankings, and genocidal propaganda.¹⁶²

Second, the conception of authority at issue here remains located primarily in individual speakers, and, as I have argued throughout this dissertation, that fails to do justice to the force of subordinating speech. In other words, the connection between the speaker and the broader norms of oppression falls from view. Yet it is often this connection that permits the speaker to make a subordinating speech act. With that concern in mind, it is useful to turn to McGowan's alternative account.

¹⁶² In making this criticism, I acknowledge that Langton has available the response that her target of analysis is itself narrow. In other words, she (perhaps) is not making any claims about the large swath of subordinating speech acts that are beyond her (current) purview.

4.2. *The No-Authority Alternative*

In this section, I consider an alternative framework for thinking about subordinating speech as presented by Mary Kate McGowan in a series of articles. This alternative is worth considering here for a few reasons. Most significantly because it eschews the necessity for speaker authority in oppressive speech. And since it was the reliance on speaker authority that poses problems for other models—namely, the Authority Problem considered in Chapter Two—the idea that discarding this concept might liberate the framework is compelling.

However, as I'll argue below, this framework comes with significant costs. Most significantly for my purposes is that it flattens the terrain of subordinating speech. It does so by viewing the capacity to alter permissibility facts (that oppress) as the hallmark of oppressive speech. That is, McGowan sees the Austinean illocution of exercitives—which set and change facts about what is permissible and impermissible—as the principal, apparently singular type of oppressive speech act. In this way, and unlike Langton, McGowan does seem to be making claims about subordinating speech as such.¹⁶³ For this reason, this view stands in stark contrast to my own approach.

Beyond this difference McGowan's account also has further, internal problems. Most notable it elides an important distinction between *enacting* new norms and *engaging* pre-existing norms. In this way, it is not unlike Maitra's account discussed in Chapter Two, where both seem to mischaracterize the relationship between individual subordinating speech acts and the broader systems of oppression that gives these speech acts their force. Before developing these

¹⁶³ For example, consider the following: “Since oppressive speech enacts permissibility facts (that oppress), speech must be exercitive in order for it to be an act of oppression. In other words, in order for a particular utterance to be an act of oppression, that utterance must have exercitive (illocutionary) force” (McGowan 2009, 392).

criticisms, I'll first explain McGowan's approach to oppressive speech, focusing on its development of the exercitive.

4.2.1. *Standard exercitives*

Austinean exercitives—which McGowan also calls “standard exercitives”—are speech acts that enact policies and rules, issue commands and directives, and make some actions permissible or impermissible. They are authoritative speech acts, where one of their felicity conditions is that the speaker possess the appropriate authority over the domain in question. Without this authority, the resulting speech act is “unhappy,” or a misfire.

For example, if a college professor says, “your papers are due on Friday,” they make it the case that students must turn in their papers by that time or receive a penalty. It is an exercitive speech act that sets what is permissible and impermissible for some course of action. If, instead, one of the students had said, “I declare that our papers are due Monday, everyone,” no change in permissibility facts would follow, and this is straightforwardly because the student lacks the authority that the professor possesses to do this action with their words.

Notice also how exercitives also contain the new rule enacted as the content of the utterance itself. Taken together, then, standard exercitives (a) require the speaker to hold an authoritative position, and (b) include the norm-shaping command as part of their explicit content. Before moving onto McGowan's adaptations of this concept, it's worth pausing to note how this speech act does capture at least some instances of subordinating speech. Notably, some of the ‘paradigm’ examples fit this model well, such the South African legislator case, the Alabama slave law, and instances of “White's Only” signs posted in the Jim Crow South. Each

case involves an authoritative speaker enacting a policy—an oppressive policy, no doubt—via a speech act whose content matches the policy being enacted.

Though this model fits some cases, it is nonetheless inapplicable for the bulk of subordinating speech. Much of what we intuitively take to be oppressive speech still requires elaboration. As Ramsey puts it, “the main problem with the [standard] model is that much oppressive speech is not obviously exercitive” (2013, 38–39), at least not so long as we stick to standard exercitives.

4.2.2. *Covert exercitives*¹⁶⁴

Because of the limitations of the standard exercitive model, McGowan considers a different type of norm-shaping; cases where no explicit authority is present and where the semantic content of the utterance and the rule enacted is less clear cut. But because they still enact permissibility facts,¹⁶⁵ she calls these utterances “covert exercitives” (2009; 2018).

What distinguishes covert exercitives from standard exercitives is their ability to enact permissibility facts without requiring (a) the speaker to hold a position of authority or (b) containing the content of the rule-change as part of the content of the utterance.¹⁶⁶ Instead, covert exercitive enact changes to what is permissible by triggering the rules of a norm-governed

¹⁶⁴ I should note that McGowan sometimes uses the concept of ‘conversational exercitives’ to bridge the gap from standard exercitives to covert exercitives, in part by drawing explicitly on the idea of accommodation in conversation, as explained by Lewis (1979). It is with this idea in tow that McGowan then takes her view beyond the norms of conversations, to include wider practices—like oppression.

¹⁶⁵ McGowan writes most often of ‘permissibility facts,’ and ‘rules,’ and is most concerned with how these are at work in ‘norm-governed practices,’ like conversations, games, or systems of oppression. I will use sometimes follow her terminology, and sometimes use the language of norms, as each term seems equivalent in the sense required.

¹⁶⁶ McGowan’s understanding of the contrast between ‘standard’ and ‘covert’ exercitives seems to shift slightly throughout her work. For example, in her (2012) she marks the distinction almost entirely on the idea that standard exercitives are an exercise of speaker authority, while noting that they may be ‘implicit’ and ‘indirect’ standard exercitives, which do not wear their exercitive force on their sleeves—that is, as part of their semantic content. In other work she builds in identity between content and enacted norm as a requirement of standard exercitives.

activity within which the utterance constitutes a move.¹⁶⁷ This is the case because “any contribution to any rule-governed activity changes what is subsequently permissible in that activity [and] most collective human activities are rule-governed in the relevant sense” (McGowan 2009, 395).¹⁶⁸ In this way, covert exercitives help explain how an utterance like, “my child’s tutor is amazing,” makes other new conversational moves now permissible, such as asking for the name of the tutor, or how old their child is.¹⁶⁹

So far, this has simply touched on the norms that guide conversation, but more is at play. Notice that some speech acts are part of activities that go beyond the scope of that particular conversation. A wedding invitation, for example, triggers norms not only related to conversation (e.g., once being invited, a response is now called for) but also the norms of the relationship within which it occurs (e.g., do the engaged really expect you to fly to Aruba for this?), as well as the norms of etiquette (e.g., is a gift now called for?), and likely much more besides. These are the broader activities that are also at play in a simple utterance, but each has its own unique norms that govern appropriate moves going forward. It is by being a move in another normative activity that speech “triggers the rules of that activity and thereby enacts facts about what is subsequently permissible in that activity” (McGowan 2009, 396).¹⁷⁰

¹⁶⁷ “When speech constitutes a move in a norm-governed activity, it has exercitive force in virtue of enacting new permissibility facts for the activity in which it is a move” (McGowan 2012, 134).

¹⁶⁸ McGowan explains her conception of ‘rule-governed activity’ as follows: “By rule-governed, I mean any activity governed by norms. The ‘rules’ in question need not be explicit, formal, exceptionless or even consciously recognized. If at least some behaviours (as contributions to the activity in question) would count as out of bounds or otherwise inappropriate (as contributions to the activity in question) then that activity is rule-governed in the relevant sense. Conversations, dancing, playing music, walking, chess, checkers, and baseball are all rule-governed in the appropriate sense.” (2009, 395).

¹⁶⁹ Conversely, it makes it impermissible for someone to respond with: “So, do you have any kids?” Impermissible not in the sense that it is impossible for this to be said, but impermissible in the sense that it goes against the norms that constitute conversation, including the near-seamless accommodation of presuppositions.

¹⁷⁰ McGowan further clarifies what counts as a move: “A *move* in a rule-governed activity is a contribution to, and thus a component of, that activity” (2009, 395). So not every action that occurs within a rule-governed activity is itself a move. As she notes, when a baseball player scratches their nose, it has no effect on what is subsequently permissible within the game.

The final pieces to add to this story are to note (1) that oppression itself is a rule-governed or normative activity, and (2) that speech acts can be contributions to this activity. McGowan supports the former point by noting that “oppression ranks people according to their membership in socially marked groups and [...] involves treating persons in some categories differently than persons in other categories” (209, 396). A sign posted with the words “Whites Only” would be a clear example of linguistic contribution to oppression, lending support to this second claim. The sign is a (linguistic) contribution to oppression, in part because it enacts the discriminatory rule that non-white people are not permitted in the establishment.

But McGowan wants to examine much more besides explicit directives like “Whites Only,” which functions like the authoritative—and thus, standard—exercitive. To illustrate the potential of the covert exercitive model, she provides the following example:

Steve and John are co-workers at a factory with very few female employees. The following exchange takes place in the factory’s employee lounge:

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

Steve: I banged the bitch.

John: [smiling] She got a sistuh? (McGowan 2009, 399)

Here, a few moves occur simultaneously. Steve’s utterance is a conversational move, making one woman salient in the conversation, allowing John to refer to her with ‘She.’ But it is also, McGowan claims, “a verbal means of mistreating women.” By this she means that “it is a contribution to, and thus a component of, a system of gender oppression. As a result, it covertly enacts permissibility facts [...] in this system of gender oppression” (2009, 399). What permissibility facts are these? As a hypothesis, McGowan suggests this speech act “makes it

acceptable, in this immediate environment and at this time, to degrade women. [Steve's] utterance makes women second-class citizens (locally and for the time being)" (2009, 400).¹⁷¹

Taking all this together, we're led to the conclusion that the utterance constitutes oppression because it enacts (local) norms that discriminate and subordinate women. It is therefore a speech act with exercitive force, and a contribution to an oppressive institution. And through this we're led to the idea that, on this view, speakers "need not possess the positional authority required of standard exercitives because the [broader norms] of the oppressive practice supply the needed authority" (Ramsey 2013, 47).

Speaker authority thus falls out of the picture as a requirement for the type of speech act that constitutes oppression. "The power to enact the permissibility facts resides," McGowan claims, "not in the speaker (as with standard exercitives), but in the rule-governed activity in question. When one performs a covert exercitive, one enacts permissibility facts, not by exercising one's own authority or power, but by *triggering the rules of the system*" (2009, 402 emphasis added). She thus has a blunt answer to the Authority Problem that vexes others. "It matters little," she claims:

whether we say that conversational exercitives are authoritative (and that all conversational participants have the requisite authority) or that they are not (since no peculiar authority is required). Either way, the authority condition is either met or inapplicable and thus poses no challenge. (2003, 180)

¹⁷¹ McGowan continues: "If Steve's utterance does this, then it is akin to a sign reading: 'It is hereby permissible, in this local environment and at this time, to treat women as second-class citizens.' Such a sign would surely be an act of gender oppression. The hypothesis suggested here is that, perhaps, Steve's utterance is too" (2009, 400).

Therefore, the power to enact new permissibility facts is not bound up in speaker authority.¹⁷²

And since it is through enacting oppressive permissibility facts that McGowan defines oppressive speech, it is then as covert exercitives that speech oppresses.

4.2.3. *Criticism of the covert exercitive model*

Having explained McGowan's account of oppressive speech, I will now argue that it is an inadequate model for the bulk of subordinating speech. This matters because it is—or ought to be—a constraint on accounts of subordinating speech that they be able to explain a wide variety of examples as instances of subordinating speech. McGowan's model, I argue, either lumps assorted acts into a single type, or excludes recognizable cases.

John Michael Ramsey thoroughly explores McGowan's model in his “How We Do Bad Things with Words” (2013, esp. Ch. 1). Ramsey argues that speech acts like those McGowan draws our attention to *engage existing* oppressive norms, rather than *enact new* ones (2013, 48). That is, the main examples McGowan uses to motivate her account “blur the distinction between engaging rules and enacting rules” (Ramsey 2013, 40). As he notes, oppressive speech often succeeds as a contribution to a conversation and to wider practices of subordinating *because* the norms are already well-established and accepted within a community. To use an example Ramsey draws on, when in 2012 Rand Paul made a (bad) joke saying he “didn't think [President Obama's] views on marriage could get any gayer,” his utterances does not make it newly permissible to disparage gay people. The ‘joke’ works in that context—if it does at all—because

¹⁷² “Since our words often are contributions to such activities (e.g., conversations, games or systems of oppression), they are often covertly exercitive. Consequently, our utterances routinely change what is permissible for those around us” (McGowan 2009, 406).

his audience at the Iowa Faith & Freedom event already accept that norm (Ramsey 2013, 50). Or at least we can assume so much based off their raucous laughter.¹⁷³ Consequently, it's a misrepresentation to characterize the oppressive force of Rand Paul's comment as one that enacts a new norm concerning homophobia. We do better to acknowledge how his comment *engages* this existing norm, and in doing so *contributes* to oppression overall.

To *engage* a norm is different than to enact one. Recall the motivating example of a (standard) exercitive is something like a parent saying, "your bedtime is 9pm," or "no games after dinner." If it already was a standing rule that the child's bedtime was 9pm or that games were not permitted after dinner, then it's hard to see how the parent's utterance is exercitive. The fact that the norm is already established undermines the idea that a new rule is being uttered into existence. And this makes sense, as we routinely engage existing norms. While sometimes new rules are enacted by fiat, "for the most part social practices [...] and the norms constituting these practices are performed and normalized in such a way that creates certain, standing permissibility facts" (Ramsey 2013, 51). When a person uses subordinating speech, they engage these standing permissibility facts, and in doing so express a commitment to these norms. This is, at least intuitively, a contrast to the practice of enacting a norm.

We see this contrast again in McGowan's example of 'Whites Only' signs. Given that in this context, this law was widespread and widely known, it makes better sense to say the sign *describes* or *expresses* the rule, rather than enacts it. While the person who posts the sign likely endorses the policy, it's unclear to what extent they are best described as the one who makes it the case. And this makes sense, since in most cases—e.g., unless they are owner with the power

¹⁷³ For video of the original comment, see Lavender (2012).

to make discriminatory policy stick in this jurisdiction—they lack the authority to make new policy of this sort.

The covert exercitive model is supposed to explain how oppressive speech constitutes oppression. According to the model, speech is oppressive when it enacts oppressive permissibility facts. And this occurs when a speaker subtly alters the existing norms that are salient in a given context. A sexist remark like Steve's in the example above makes it the case that sexist comments—and behavior too, maybe—are now acceptable, at least in the here and now, and so long as the comment doesn't receive substantive challenge from anyone. To be sure, McGowan clarifies that the norms being enacted are often highly local and may evaporate a short time later. But what this model has difficulty acknowledging is how antecedently existing norms are (often) already established in that context. While it makes some sense to see these speech acts as further entrenching background norms and making them vivid in a new, more specific context, this analysis seems to collect too much under one concept.

Where McGowan's model makes the most sense is in diagnosing how initial utterances can make certain norms more salient—creating new permissions in the here and now—but it is less effective in representing subsequent utterances as doing the same thing. That is, it is odd to say that John's utterance, which also engages sexist norms, is also enacting these same norms once again. But his utterance is also sexist and, we can assume, also a subordinating speech act. According to McGowan, however, this would mean that it ought to be covertly exercitive. While it surely contributes to the ongoing sexist norms that Steve's utterance first made salient, to say it enacts these norms seems to suggest that he is working from scratch, rather than building upon

an existing edifice.¹⁷⁴ If we were to imagine Steve and John bantering in this way, back and forth, for ten minutes, it would be strange to characterize the last utterance as newly enacting sexist norms. But this is what (covert) exercitives do. The model therefore loses its appeal once we consider subordinating speech in contexts where oppressive norms are already established and have a strong grip on social relations—that is to say, nearly everywhere.

So, one problem with the covert exercitive model is how it represents utterances as *enacting* new norms when they more properly are described as *engaging* or *maintaining* these norms. In this way the covert exercitive model “does not adequately explain how speech acquires oppressive force” (Ramsey 2013, 53).

A related problem is how the model fails to acknowledge the variety of subordinating speech acts. It does so by (1) mischaracterizing them all as exercitive, when at least some are better seen as engaging existing rules. But it also does so by (2) mischaracterizing the different speech capacities of different speakers.¹⁷⁵ That is, by dispensing with the need for speaker authority, and instead claiming that all speakers are suitably positioned to enact oppressive norms, it fails to note that different speakers have different speech capacities in front of them. In doing so, it incorrectly sees all subordinating speakers as doing roughly the same thing—enacting oppressive permissibility facts. In this way it mirrors Langton’s ‘approaching authoritative speech’ model, discussed above. But unlike Langton, McGowan takes herself to be analyzing oppressive speech in general.

¹⁷⁴ In this way again, McGowan’s approach has similarities to Maitra’s discussed in Chapter Two. This is no surprise as they both are using an Austinian framework to model oppressive speech. In this way, they both tend to focus on isolated, individualized speech acts, while either bracketing or ignoring other relevant features of subordinating speech. In other words, both accounts carefully attend to the pragmatics of singular speech acts, but this focus obscures the specifics of how the broader, structural features of oppressive systems make subordinating speech acts unlike other speech acts.

¹⁷⁵ By ‘speech capacities’ I mean the different speech acts available, at a given time, to speakers occupying different social locations. See Ayala (2016), though she uses the term ‘speech affordances.’

For this reason, it's useful to see why her account goes wrong. In short, McGowan sees *any* conversational contribution as exercitive.¹⁷⁶ In doing so, she loses sight of what makes exercitives an interesting category of speech acts. Since, if everything is an exercitive, then there is nothing notable about this category of illocutions, and labeling speech acts as exercitive loses its significance.¹⁷⁷ While McGowan is correct to note that all contributions to a conversation alter what is permissible to occur following that contribution, this fact alone is trivial rather than the hallmark of oppressive speech. We notice this most explicitly when we take note of exactly what McGowan pushes aside: differences in speaker authority and entitlements.

While all speech has the potential to alter the normative landscape it occurs in, not all speakers have equal speech capacity (Ayala 2016). This idea is familiar and underwrites the common feminist observation that women's comments are routinely ignored,¹⁷⁸ and only taken up when uttered by a man. In other words, people in different socially-relevant positions are able to do different things with their words. This point was significant in the previous chapter.

To be sure, McGowan does acknowledge that "there is nevertheless a sense in which covert exercitives require that the speaker have *a certain status*" (2009, 402 emphasis added). But, on her view, this is a status we all share, as being participants in the rule-governed activities of conversations, etiquette, oppression, etc.¹⁷⁹ Of course, this leaves open whether we are all

¹⁷⁶ Consider: "I argue that any conversational contribution invoking a rule of accommodation changes the bounds of conversational permissibility and is therefore an exercitive speech act" (McGowan 2004, 169).

¹⁷⁷ Herbert and Kukla (2016, 578) note how it is only in "a rather forced sense" that we could say that many speech acts enact new norms or rules—in their case: ingrouping and outgrouping.

¹⁷⁸ Social epistemologists have discussed this under the rubric of "testimonial injustice" (Dotson 2011) and "epistemic injustice" (Fricker 2007). Tirrell (2018b) offers an analysis of this problem using the tools of a philosophy of language approach, as does Kukla (2014).

¹⁷⁹ McGowan elaborates as follows: "Of course, there is a trivial sense in which such speakers have such authority. In virtue of being a participant in the rule-governed activity in question, each participant has authority over how that rule-governed activity evolves in the case at hand. The baseball player may not have authority over the rules of baseball as such but he does have authority over how this particular baseball game is played. He has such authority in virtue of being a participant in the game in question" (2009, 402n).

equal participants in specific normative activities. In baseball, the catcher can't just grab a bat and hit a home run. Differences in status delineate the moves available to each participant, and I argued above that something similar occurs within oppression's implicit rules.

But McGowan's account of how individual oppressive speech acts connect to broader systems of oppression is unhelpful, mostly because it is so binary. Either a speaker is saying something that participates in the system of racism, or they are not. It therefore doesn't admit of the nuances I believe are necessary to draw out. What's needed is a fuller depiction of how individual speech acts draw from and contribute to systems of oppression and domination. McGowan is clearly open to this understanding, as she says that "it seems reasonable to suppose that who can make which sort of move (in a rule-governed activity) is sensitive to *one's position* in that activity" (2009, 403 emphasis added).¹⁸⁰ However, her exploration of this idea is brief, and is focused more on the idea of *undoing* certain norms. So, what we're left with on the covert exercitive model is a limited understanding of what subordinating speakers *do* with their words, and how different speakers have differing potential. This offers a blunt connection between the individual acts and oppressive norms, which leaves no room to identify different kinds of oppressive speech acts that are importantly distinct.

I believe that we must go beyond the models offered by both Langton and McGowan and consider the diversity of subordinating speech unencumbered by the limits of only one speech act type.

¹⁸⁰ She adds: "it may be that men are in a privileged position (with respect to the ability to make various sorts of moves) in a system of gender oppression. If this is correct, then it may be especially difficult for women to reverse certain sorts of s-rules enacted by men" (2009, 403).

5. CONCLUSION

My goal in this chapter was to demonstrate existence of multiple types of subordinating speech acts. I wanted to show that there are many different sorts of speech acts that oppress, and these do so in different ways, drawing on different entitlement conditions on the speaker's side of things, and plausibly generating different outputs on the target's side. This should be an obvious and trivial fact. However, as I argued in section 4, dominant models of oppressive speech have a hard time recognizing it. Langton's "approximating authority" model depicts authority as primarily a feature of individual speakers, and I argue that this presents a monolithic understanding of subordinating authority. On the other hand, Mary Kate McGowan's "no-authority" model offers an intriguing account of how these background norms supply oppressive force in lots of circumstances. But, by only conceiving of a blunt connection between speakers and these norms, it flattens the terrain and does not allow for the nuances that animate varieties of subordinating speech.

Instead, I opted for a bottom-up approach, and endeavored to catalogue a diversity of subordinating speech acts. From this grouping I pulled out a variety of distinct types of subordinating speaker authority and entitlements, along with an account of how to delineate differences in pragmatic output. Using this approach, I identified a type of subordinating speaker authority that is, in an important sense, collectively-sustained. These *collective speech acts*, as I call them, reveal instances of subordinating speech where the entitlement—indeed, subordinating authority—speakers draw on is not reducible to an individual position. Instead, these speech acts draw power and mutual support from taking place within a group practice. I believe this conception helpfully opens new avenues for discussing real cases of subordinating speech, and in the following chapter I apply this idea to the under-analyzed case of online harassment.

CHAPTER FOUR:
WHOM DO YOU SPEAK FOR? AND HOW?
ONLINE ABUSE AND COLLECTIVE SPEECH ACTS

“Internet trolls, predominantly anonymous posters, realized they could work together to try to destroy the lives of people who disagreed with them.”
- Ian Sherr and Erin Carson (2017)

“I don't think Facebook changed the way people are in any negative or positive way. It just showed how people are.”
- Corey Barrow, founder of Mad World News (quoted in Roose 2018.)

1. INTRODUCTION

The main topic of this chapter is online abuse. I have two goals in directing our attention here. First, I want to show that this is a serious, but neglected, area of subordinating speech, and that social philosophers of language have good reason to pay more attention to the harms of online discourse. Second, I will argue that accounting for the realities of online abuse shows that speaker authority is dynamic and emergent, and often depends on the community in more ways than standard theories of licensing and accommodation are equipped to explain. I argue that much of online abuse is best understood as a type of *collective subordinating speech act*. Overall my hope is to show that attention to online subordinating speech is useful for illuminating both the harmfulness of that phenomena itself, but also for clarifying features of IRL (in real life) speech that regularly go under-emphasized in the existing literature.

It's not controversial to say that a lot of subordinating speech now occurs online. Yet much of the philosophical work in this area has focused on offline life. This immediately raises several questions for philosophers. Can current accounts of oppressive speech adequately capture digital hate? How does the anonymity of online harassers contribute to the force of their speech?

To answer these questions, I address the varied roles of *identity* in online abuse—for both its targets and its speakers—and argue that the combination of anonymity and shared language offer online abusers a path to a type of group-authority that lends more power to their speech than they might first appear to have.

By exploring the management and attribution of identity in online abuse, I argue that online abuse is best understood as a type of *collective subordinating speech*, rather than speech acts originating from individual speakers. Even where there is an identifiable single speaker and they address a single person, because the speech is semi-public¹⁸¹ and because these tropes and messages are used widely and repeatedly, we miss something important when we focus too closely on any one interaction apart from its wider context. Online abusive speech is speech where the ‘speaker’ of these messages is better conceived as a collective.

To make this argument, I return to the model that claims speakers can come to gain the authority they lacked prior to speaking through processes like *licensing* and *accommodation*. The basic idea is that while the hate speaker can lack the necessary authority to subordinate before they make their utterance, because of the silence of bystanders, the audience fails to block the speaker’s assertion, imbuing it with subordinating force. This approach has proven quite popular, and yet I argue that it fails to explain the dynamics of online abuse, and that this failure reveals a more widespread tension in the concept.

I begin in section 2 I by noting a few ways online subordinating speech is importantly different than IRL subordinating speech, especially with regards to online propaganda. It’s useful to begin here, as this is the area that has received the most philosophical attention (that I am

¹⁸¹ As Regina Rini (2018b) has remarked, the design of social media platforms themselves promote the manner in which speakers speak to a larger audience even when engaging in direct interpersonal exchange.

aware of). Next, in section 3, I describe some the key features of online abuse, and I explain how these features pose a problem for some existing accounts of the force of subordinating speech, particularly around the notion of accommodation. This leads me develop an alternate conception of the subordinating authority at work in online abuse, and section 4 is devoted to developing this idea. By focusing on (1) the role of anonymity and (2) the use of shared language, I claim that much of online abuse is best understood as a *collective speech act*. As I argue, accounting for the realities of online shows that speaker authority is dynamic and emergent, and often depends on the wider community in more ways than simple accommodation.

I argue throughout that much of online abuse challenges existing accounts of subordinating speech, and it therefore represents in some ways a distinct phenomenon. At the same time, though, I believe that it can also shed some light on IRL subordinating speech. That is, I aim to show throughout how online speech makes explicit many features that it shares with IRL hate speech, but which tend to be ignored or de-emphasized in existing accounts. Despite the internet offering bigots and abusers new tools and strategies, for the victims, the experience of being targeted of such abuse is remarkably similar in some key ways. The examination of online abuse therefore helps to reveal key features of subordinating speech across mediums, and I summarize these in my conclusion.

2. BACKGROUND: THE INTERNET AND ITS PROBLEMS

As the work cited in this dissertation shows, philosophers have been writing about subordinating speech for a few decades now. And while the internet have been around almost as long, much of the philosophical work on hate speech, propaganda, and subordinating speech in

general has focused on offline life.¹⁸² In-person hate speech like what you might see in public spaces (Maitra 2012; Langton 2018a; 2018b; McGowan 2012), propaganda as it is disseminated in print or on the radio (Stanley 2015; Tirrell 2012; Smith 2012), and more recently microaggressions as they occur in say, a workplace or college classroom (Liebow 2017; Rini 2018; Saul 2018b), are the main examples. This has remained the case even as more and more of our lives have migrated online.

These ‘real world’ phenomena are still worthy of philosophical analysis. But online speech raises many new issues for social philosophers and philosophers of language, and these are only beginning to be explored. Fundamental questions like who or what should count as a ‘speaker,’ or how retweets, ‘likes,’ ‘favs,’ and emojis fit into an account of utterances, all need to be re-examined. As does my current question: how has the internet changed subordinating speech?

How much the internet has changed the way we use language is subject to disagreement. One tempting answer is to say that the internet has *radically changed everything!* And there is at least some truth to this, even for more familiar speech acts, like apologies, retractions, quotes, etc.¹⁸³ We may think this is especially likely of subordinating speech—in part, since if you give the oppressor a new tool, usually they’ll use it to oppress. And the internet offers a lot of new tools.

¹⁸² For a quick and non-decisive example, consider that the index for the (2012) anthology *Speech & Harm: Controversies over Free speech*, edited by Ishani Maitra and Mary Kate McGowan has no entries for the terms ‘internet,’ ‘website,’ ‘online,’ or other specifically online communication mediums. A few noteworthy exceptions include: Barney (2016), Cherry (2015a), Frost-Arnold (2016), Herbert (2018b), Levmore (2010), Manne (2017), Nguyen (2018), Nussbaum (2010), Rini (2018b), and Technau (2018).

¹⁸³ For one example, see Norlock (2017a) on apologies.

2.1. Epistemic Problems

Take propaganda. One initial thought might be that all that the internet has done is made it easier to spread hateful propaganda to more people, more quickly.¹⁸⁴ And if this were all it did, that would be problem enough. However, recent analysis suggests things are more complex, and more insidious, than that. Writing for *Wired*, Issie Lapowsky notes that “[t]he long past of propaganda blended with the communication channels of the present and future form a toxic mix” (2017).¹⁸⁵

The reach and speed of the internet is but one concern. These, along with other new technologies that have their home online point to a difference in *kind*, rather than only degree. In addition to potentially reaching millions of people in mere seconds, photo-, video-, and audio-editing tools offer new tools in the propagandist's toolbox. These have profound effects on the epistemic dimension of propaganda, with the potential to make otherwise unbelievable things frustratingly credible. For example, researchers at Stanford University have developed software capable of “manipulat[ing] video footage of public figures to allow a second person to put words in their mouth—in real time” (Solon 2017; and see BBC News 2017 for a video demonstration). More generally, “deepfakes,” that is, videos created using machine learning to create the illusion that someone has said or done something they never did, are becoming more widely available as the technology improves.¹⁸⁶

¹⁸⁴ See Ellul (1965) for an argument as to why propaganda will “leave no part of the intellectual or emotional life alone. [and will] surround [one] on all sides [...] by all possible routes” (quoted in Lederer 1995, 133).

¹⁸⁵ On this toxic potential, two philosophers quoted in the *Wired* article seem to agree: “‘I think this is real dangerous shit,’ says David Livingstone Smith, a professor of philosophy at the University of New England, who specializes in the history of dehumanization and who authored a book on the topic called *Less Than Human*. [...] ‘This is scary shit,’ echoes Jason Stanley, a professor at Yale and author of the book *How Propaganda Works*, whose father fled Nazi Germany in 1939 (Lapowsky 2017).

¹⁸⁶ As Rini (2019) describes them: “Deepfakes are fabricated video or audio recordings created through machine learning technology. A computer program uses a large data set of real recordings to build a model of the facial/vocal

The above capture cases where someone is willing to put the effort in to create an elaborate fake; often, this is unneeded. While the norms here are still evolving, many of us are still disposed to think that images, and especially video, give us a less-filtered version of reality than mere testimony. And this has clear repercussions for the type of propaganda produced. As Lapowsky notes:

It's not simply cartoons and phony headlines filling people's minds. Doctored photos and misrepresentations of real footage, like the video the President shared [of supposed 'Muslim violence'], are a dangerous new development in the history of propaganda, experts say. "Everyone knows caricatures exaggerate," says Claudia Koonz, a historian at Duke University and author of *The Nazi Conscience*, "but gullible viewers, including probably Trump, see videos as reality." (Lapowsky 2017)

In other words, sometimes even small changes to photos and videos—like the simple technique of slowing video down to make speech sound slurred, as occurred for House Majority Leader Nancy Pelosi (see Harwell 2019)—can be just as effective as an elaborate fake.¹⁸⁷

The epistemic problems online go deeper when we consider how testimony functions on social media. Even as testimony, it is, as Regina Rini (2017) sees it, a "bent" form of testimony whose features exacerbate the pre-existing problems posed by something like fake news, which is plausibly an old sub-branch of propaganda. For Rini, fake news is not limited to online communications, but there is, as she says, "a strong contingent relationship between fake news and social media" (2017, 45), making the one ripe for the other. As she says:

I suspect that the two bent features of social media testimony are related to one another. Perhaps people are less inclined to subject ridiculous stories to scrutiny *because* we have unstable testimonial norms on social media. A friend posts a ridiculous story, without comment, and *maybe* they don't really mean it. But then other friends 'like' the story, or

characteristics of a person, then superimposes this onto recordings of another person. The effect is an apparent recording of a well-known person doing or saying something they never did."

¹⁸⁷ It is worth acknowledging that a common and effective technique to suppress dissident views is simply to flood the information channels with irrelevant and sometimes contradictory content. The thought being that it is easier to drown out 'unwanted' information than to censure it. Zeynep Tufekci (2017) describes this aspect in relation to how networked communication has changed things for both revolutionary social movements and repressive regimes. The philosophy Michael P. Lynch (2016) makes a similar point.

comment with earnest revulsion, or share it themselves. Each of these individual communicative acts involves some ambiguity in the speaker's testimonial intentions. But, when all appear summed together, this ambiguity seems to wash away. Perhaps the implicit thought is like this: could it really be that *all* these people aren't *really* testifying to this? A thought like that might overwhelm ordinary skepticism about ridiculous testimony. (2017, 49)¹⁸⁸

Rini's analysis shows how fake news can spread originally, given the unstable norms of social media (as they relate to testimony, in particular). This means that little to no malicious intent is needed. But, as Zeynep Tufekci (2017, 241) notes, "social media's business model financed by ads paid out based on number of pageviews makes it not just possible but even financially lucrative to spread misinformation, propaganda, or distorted partisan content that can go viral in algorithmically entrenched echo chambers." And, of course, there is little incentive for the private companies that own and operate these social network platforms to eliminate this type of content.

In sum, the internet—and social media in particular—not only amplify the reach and speed of propaganda, but also increase its credibility in several distinctive ways.

2.2. *Non-Epistemic Problems*

All the above cases all involve deception, though perhaps sometimes unwitting. But another development flips the direction of fit and instead aims to make the world as the propagandist wants it, and then captures that for distribution. *ProPublica* reported on a new white-supremacist group called RAM (the Rise Above Movement) responsible for some of the violent assaults in Charlottesville, Virginia, in August 2017. They note how RAM members have

¹⁸⁸ Chloé Bakalar (2018) offers a different analysis of the epistemic features of social media testimony that focuses on the problems posed by Automated Content Moderation Algorithms (ACMAs).

teamed up with Vincent James Foxx, “a 31-year-old video blogger and livestreamer with a fondness for white supremacists and radical right-wing politics” who operates as their unofficial propagandist (Thompson, Winston, and BondGraham 2017). Writing about one violent rally in Berkeley, California, they note that:

Foxx wasn’t just documenting the violence at the rally—*he was inciting it*. On video Foxx posted to YouTube, he can be heard repeatedly encouraging RAM members and others to assault people. “Get that fucking cuck!” he screamed when a RAM member and four or five other men grabbed a counter-protester and began beating him. “Charge!” Foxx yelled as a mob of right-wingers went on the offensive. (Thompson, Winston, and BondGraham 2017, emphasis added)

Here, rather than affecting the reach, speed, or credibility of propaganda, social media is itself providing the impetus for physical violence—the ability to document violence for the purposes of spreading propaganda online served as motivating cause of his incitement. I call attention to this since violence is usually thought of as the downstream ‘mentally mediated’ consequence of propaganda. But here, rather than propaganda begetting violence, we have violence begetting propaganda. At least in a sense.

Indeed, the crossover between online hate and real-life violence is hard to deny. After the New Zealand Mosque shootings, in a piece titled, “Mass Shootings Have Become a Sickening Meme,” *New York Times* writer Charlie Warzel (2019) wrote:

it’s becoming increasingly difficult to ignore how online hatred and message board screeds are bleeding into the physical world—and how social platforms can act as an accelerant for terroristic behavior. The internet, it seems, has imprinted itself on modern hate crimes, giving its most unstable residents a theater for unspeakable acts—and an amplification system for an ideology of white supremacy that only recently was relegated to the shadows.

And the connections aren’t only there for one-off events, but also for broader cultural shifts. For example, *New York Times* reporter Kevin Roose speculates that the violent events in Charlottesville in 2017 wouldn’t have occurred had Discord—a chat app meant for gamers but

also a home for the ‘alt-right’—and other platforms shut them down earlier (Barbaro 2017). And some take the combined presence of Twitter and Facebook to have been necessary conditions for President Trump’s 2016 election win (for one example, see Sherr and Carson 2017).¹⁸⁹ Whether or not this is true, it has led to much soul-searching among the tech giants, as demonstrated by Facebook’s recent “Hard Questions” series (see Harbath 2018).

In sum, these issues lead many to think, reasonably, that online hate is substantially different from what we’ve seen before, and it changes the game compared to more traditional forms of propaganda. In short, there is therefore good reason to think the rise of the internet—and social media in particular—make online hate speech different from what we’ve seen before and is a cause for alarm. The worldwide concern for fake news demonstrates this well. We are only now beginning to see the effects of these developments, and some analyses are quite damning. Violence arguably caused by online propaganda and fake news has been reported in countries ranging from the US, Myanmar, Germany, India, Canada, and others.¹⁹⁰

It is noteworthy, however, that the bulk of this analysis focusses on subordinating speech as it functions in its propagandistic mode—as outreach—rather than cases where it is directly targeting and harming particular individuals.¹⁹¹ This has led to discussions of the ‘potential’ harms of online hate and have focused on abstract values ‘democracy’ as its main victim. But this ignores those who have *already* been victimized by online hate and minimizes the harm they’ve experienced—this is no surprise, as most of *these* victims are not white men. In what

¹⁸⁹ That is, how a fringe online movement comprised of angry gamers and angry racists—what some have called ‘the Chanterculture’ after the notorious online forum 4chan—played a key role in elevating Trump to the presidency.

¹⁹⁰ Indeed, as I mentioned in Chapter One, the Quebec City Mosque shooter was partly inspired by online extremist content, and then the final straw was a tweet by Prime Minister Trudeau. See Cecco (2018) for details.

¹⁹¹ Note that a single speech act can play both roles at once, though. I discussed these differences and where they overlap in the previous chapter.

follows, I will examine online abuse as a topic worthy of serious philosophical investigation. I consider it as a structural problem, but I want to draw attention to how it is experienced by those targeted by online abuse.

3. ONLINE ABUSE

To being, we need a better idea of what online abuse includes.¹⁹² Media studies professor Emma Jane (2014) articulates the breadth of the problem well in her (aptly titled) paper, “‘Your a Ugly, Whorish, Slut’: Understanding E-bile.” Jane coins the term ‘e-bile’ to capture what she describes as the “extravagant invective, the sexualized threats of violence, and the recreational nastiness that have come to constitute a dominant tenor of Internet discourse” (532).¹⁹³ As Jane’s investigation shows, ‘e-bile’ is found in nearly all corners of the internet and displays impressive flexibility in terms of functional use. She writes that:

hyperbolic vitriol—often involving rape and death threats—has become a lingua franca in many sectors of cyberspace. It is a commonsensical, even expected, way to, among other things: register disagreement and disapproval; test and mark the boundaries of on-line communities; compete and create; ward off boredom; prod for reaction; seek attention; and/or simply gain enjoyment. (ibid., 542)

And yet, despite being so ubiquitous, and despite being put to many uses, “the rhetorical constructs of individual e-bile texts are strikingly similar in terms of their reliance on profanity, ad hominem invective, and hyperbolic imagery of graphic—often sexualized-violence” (ibid., 533). So, ‘e-bile’ is found in all corners of the internet, is used to perform a variety of speech

¹⁹² For some recent first-person accounts that touch upon the varied features of online abuse in detail, see Koul (2017), La (2017), Quinn (2017), Valenti (2016), and West (2016). For more, see Adler (2018), Bartlett (2014), Bernstein (2015; 2018), Jeong (2018), Kakutani (2018), as well as Ronson (2015) for a vivid account of online shaming, and Norlock (2017) for a philosophical analysis of the same.

¹⁹³ The topic under discussion goes by a few names: “e-bile,” “cyber-bullying,” “online harassment,” and more. I go with “online abuse,” partly to follow internet safety activist Zoë Quinn, who suggests: “the term ‘online abuse’ is far more accurate because it perpetuates the dynamics of real-life abusive situations” (2017, 50).

acts, but at the same time is strikingly similar across these usages, with expressions of sexual violence being a prominent trope.

Notably, Janes says in many cases “e-bile appears to be a pleasurable—albeit competitive—game, in which players joust to produce the most creative venom, break the largest number of taboos, and elicit the largest emotional response in targets.” For this reason, Jane concludes, “what looks like hate speech might better be classed as “boredom speech” or “gaming speech” (ibid., 534). However, as Jane is quick to note, the effects are very serious for the targets of e-bile. She reports that “women who have been targeted by e-bile generally report [...] emotional responses ranging from feelings of irritation, anxiety, sadness, loneliness, vulnerability, and unsafeness; to feelings of distress, pain, shock, fear, terror, devastation, and violation” (536).¹⁹⁴

What Jane, along with others like Zoë Quinn (2017), notes here is that the functions and motives behind abusive rhetoric are more diffuse than might be expected. And this is true both when the utterances are directed towards ingroup members, who are likely ‘in on the joke,’ but more importantly also when directed at outgroup members, what I have been calling the targets of such speech. This leads many to dismiss online abuse as mere joking around, not worthy of much attention, or moral outrage. And sometimes this fact is taken to mean that there is simply nothing to be done about the problem.¹⁹⁵ The thought seems to be, so long as these utterances are not motivated purely by hate then they can’t and shouldn’t be thought of as on par to in-person

¹⁹⁴ Further research backs up these claims. See the Women’s Media Center Speech Project for a brief overview of relevant survey data (www.womensmediacenter.com/speech-project/research-statistics/) And as all this shows, while there is a lot of variety in terms of what motivates speakers to use invective of this sort, along with countless ways in which it is put to use, so is there variety in terms of how targets response. As Quinn reminds us, “[t]here is no one story of online abuse” (2017, 69), and this must be kept in view as we develop an analysis.

¹⁹⁵ Lindy West (2016) sums up this view as boiling down to: “The Internet’s a cesspool. That’s just the Internet. We all get rude comments. Can’t make an Internet without getting a little Internet on your Internet!”

hate speech—though those who say so much are often skeptical of the importance of IRL hate speech as well.¹⁹⁶ And since they aren't *really* engaging in hate speech, best to just 'ignore the trolls,' which is the standard advice given to victims of online abuse.¹⁹⁷

While this may appear to mark an important difference between online and IRL hate speech and is sometimes touted as such; it really isn't. And here is where I'd like to first claim that online subordinating speech helpfully reveals aspects of IRL subordinating speech that, while present, too often fall from view. What is often obvious for cases of online hate speech is obfuscated in cases of IRL hate speech, distorting our understanding of both.

Many emotions besides hate motivate hate speech. Love, even, as in the love of one's community or traditions—or say, of the white race—serve equally well as the affective underpinning of hate speech (MacKinnon 1991;¹⁹⁸ Smith 2017). As Jeremy Waldron puts it, “hatred is relevant not as the motivation of certain actions, but as a possible *effect* of certain forms of speech,” that is, what this speech aims at or is likely to incite (Waldron 2012, 35, emphasis in original).¹⁹⁹ But, too often, people see underlying hatred as the definitive mark of

¹⁹⁶ See Manne (2017, esp. Ch. 1) for a critique of this psychological view of misogyny.

¹⁹⁷ Worth pointing out that the catchall term 'troll' is not always helpful in itself. Over a few decades, it has undergone a fairly significant change in meaning, initially referring to someone who insincerely asks obvious questions for laughs, to now meaning an outright hateful speaker who regularly utters death and rape threats. See Quinn (2017, 49) for one account of the history.

¹⁹⁸ Consider: “Hatred rationalizes and impels genocide, certainly, but so do some things far colder, like self-interest, sense of superiority, or fun, and something far more banal like indifference or system. In the case of women and men, love deals at least as much death, and so does something hotter, like pleasure. The fact that pornography so often presents itself as love, indeed resembles much of what passes for it under male dominance, makes its construction as hate literature a challenging exercise in demystification, to say the least. The concept of discrimination aims not at what is felt by perpetrators or victim or what is said as such, but at what is done, including through words” (MacKinnon 1991, 808).

¹⁹⁹ He elaborates: “Many statutory definitions of what we call hate speech make the element of ‘hatred’ relevant as an aim or purpose, something that people are trying to bring about or incite. For example, the Canadian formulation [...] refers to the actions of a person ‘who, by communicating statements in any public place, incites hatred against any identifiable group.’ Or it is a matter of foreseeable effect, whether intended or not: the British formulation refers to speech that, in all the circumstances, is ‘likely to stir up hatred.’” (Waldron 2012, 35).

hate speech.²⁰⁰ And in doing so they misrepresent the variety of motives that serve to animate hate speech.

For whatever reason, though, this occurs less often with online abuse. People tend not to assume hatred is the main motivation. Instead, the deflections of its seriousness turn to the potential non-hateful motives of the speaker—or, poster, to be more accurate. This is why it is supposed to matter that they were ‘just trolling,’ or ‘being ironically racist,’ etc. It is, therefore, the recipients who are at fault for not being able to take a joke. (If they only knew what was really going on in the heads of their abusers, they wouldn’t be so upset!)

In order to address either variant, it’s necessary to recognize that subordinating speech, in both its online and IRL modes, is often directly motivated by attitudes other than hate, but this doesn’t make it less dangerous. This is a basic point, but one that deserves attention, and one that I believe the analysis of online abuse helpfully clarifies. While its motives and superficial purposes might vary—one-upping, building solidarity, expressing love for one’s own culture—a more insidious function sits just below the surface: the intimidation of outsiders in order to exclude, and the reification of existing hierarchies of domination.

By recognizing the red herring that is the speaker’s underlying psychology, and in particular the irrelevance of their (stated) motives, we’re led to put the focus back on the illocutionary act the speech performs, along with its expected perlocutionary effects. This raises the question, though, of what authority, if any, renders these speech subordinating. In the next section, I return to the philosophical literature on subordinating speech to demonstrate why it is not up to the task of assimilating online abusive speech into its apparatus.

²⁰⁰ This is unfortunate, in part, because it seems to suggest that legislation that aims to tackle hate speech looks to correct these ‘bad’ attitudes—and worries of ‘thought crimes’ slip in (Waldron 2012).

3.1. The Authority Problem: Redux

As we've seen, the authority conditions that enable subordinating speech is one topic that has received sustained philosophical analysis. And as we've also seen, many compelling answers to explain this authority have been developed, ranging from the claim that, in fact, speakers do not require any special authority to subordinate with their words (McGowan 2012), or if they do all that's needed is a type of informal authority within a given domain (Langton 1993; 2017b), to models that show how speakers can come to gain the authority they lacked prior to speaking through processes like licensing and accommodation (Maitra 2012; Langton 2017a; 2018).

This last approach has proven quite powerful, though, as I explored in Chapter Three, it is not without its inadequacies. Recall: the basic idea of licensing or accommodation is that while a speaker can lack the necessary authority to subordinate before they make their utterance, because of their silence, the audience fails to block the speaker's speech act—along its presupposition of authority—and its subordinating content is therefore successfully added to the speech situation, understood as the 'score' and/or the 'common ground.' As we saw in the last chapter, the thought is that speech acts, "including directives generally, and hate speech specifically, can acquire authority by an everyday piece of social magic: authority gets presupposed, and hearers let it go through, following a rule of accommodation" (Langton 2018).²⁰¹

Some initial roadblocks make all these accounts unsatisfying for online abusive speech. First, while an account that relies on informal conception of authority—e.g., one that picks up on parameters of privilege like race and class—has intuitive plausibility to account for the authority of IRL hate speakers, it will have a harder time explaining how an anonymous speaker, whose

²⁰¹ See Ayala and Vasilyeva (2014) for a fuller account of the specific harm that bystander silence can contribute.

only physical presence might be a cartoon avatar on the target's screen, still retains the authority that's akin to an in-person confrontation. While it is probably safe to assume the person hiding behind the Nazi anime character is white and male, convincing the police to take flesh and blood Nazis spewing hate seriously is hard enough; anime Nazis are, unfortunately, even more easily dismissed, and it seems to me this dismissal arises in part from suspicion around the speaker's power over the target.²⁰²

Moreover, according to the standard picture of licensed authority, blocking—where an audience member rejects or challenges the speaker's utterance—should be sufficient to cancel the subordinating content from being accommodated. As Langton describes it: “a hearer who blocks what is presupposed, also blocks the speech act to which the presupposition contributes. [...] That is why blocking a presupposition can make the speech act fail” (2018). And it is worth emphasizing that Langton is here referring primarily to the *illocutionary* success of a speech act, not its *perlocutionary* effects (though it can affect this too),²⁰³ and this is because blocking prevents—or rather undoes—the acquisition of authority. “A successful blocker,” she says “changes a past utterance from the unactualized way it would have been, to the way it actually is. If a speaker's presupposed authority is blocked by a hearer [...] that blocking changes the past” (2018).²⁰⁴

²⁰² This is not to say police actions are in any way a reliable indicator of speaker authority or illocutionary force. Nonetheless, I think this attitude is fairly widespread, though it does seem to be waning. Both are evident in an infamous tweet written by musician Tyler, the Creator (2012), along with some of the response it received. The tweet, posted on December 31, 2012, read: “Hahahahahahaha How The Fuck Is Cyber Bullying Real Hahahaha Nigga Just Walk Away From The Screen Like Nigga Close Your Eyes Haha.” A casual glance at the replies shows both agreement and condemnation.

²⁰³ To see both sides of this, Langton says that “Besides interfering with persuasion—with ‘perlocutionary’ success, in Austin’s terms—blocking can interfere with the speech act itself, its ‘illocutionary’ success.” And later: “Blocking prevents illocutionary accommodation, tracked by score, and perlocutionary accommodation, tracked by common ground, achieving the latter, because it achieves the former.” (Langton 2018).

²⁰⁴ As she further explains this: “Blocking can disable, rather than refute, evil speech. It can make speech misfire, to use Austin’s label for a speech act gone wrong. It offers a way of ‘undoing’ things with words (to twist his title)—

But in cases of online abuse, this does not seem to be what happens, or so I argue.²⁰⁵

Consider how in cases of online abuse, a target might receive hundreds of messages, including some that are supportive right alongside others that encourage suicide or worse. Here, there is no single, linear conversation to map a conversational score or common ground onto. This makes it is obvious, in my mind, that counter-speech standardly fails to render these situations or these speech acts non-subordinating. While it can help,²⁰⁶ it doesn't do the job of "blocking" or "cancelling" a move in a language game that philosophers sometimes attribute to it.²⁰⁷

Why is this the case? One answer emerges from considering the speech acts being performed here in more detail. As Jane notes about 'e-bile,' "the point is rarely about winning an argument via the deployment of coherent reasoning, so much as a means by which discursive volume can be increased—e-bile is utilized, in other words, to out-shout everyone else" (2014, 534). Seen in this way, it becomes clearer why more speech—blocking speech—won't work.

Recognizing that its point is not to add new content to the conversational score—content that might be contested—but instead to inundate its targets with a barrage of hurtful words and imagery²⁰⁸ shows the limits of this standard approach when the assailants number in the dozens, hundreds or even thousands. Seeing this speech for what it is thus explains the question about

and this 'undoing' has, I shall suggest, a retroactive character, which Austin himself described. It offers a ticket to a modest time machine, available to anyone willing and able to use it" (Langton 2018).

²⁰⁵ I should acknowledge that Langton is perhaps not addressing this more direct *type* of subordinating speech, as her concern is more with what she calls 'back-door' speech acts. Consider: "We are asking whether evil speech could itself misfire, could itself be disabled, when hearers exploit blocking, as counter-speech. I shall suggest a special role for blocking as a response to 'back-door' speech acts, as we might call them: low profile speech acts, enabled by presuppositions and their ilk, that tend to win by default" (2018).

²⁰⁶ As Lynne Tirrell says (about IRL speech): "Challenges tend to push the game backward—they cannot undo the move but they can revoke a license. [...] Over time, enough challenges or challenges of the right kind might kill the viability of the move, depending on how local or global the challenge becomes" (2017, 143).

²⁰⁷ Langton (2018) does list a number of 'handicaps' on blocking. Some making blocking impossible, others just making it difficult. But what she addresses here are unlike the problem I am posing here.

²⁰⁸ The important role of graphic sexual and violent imagery in online abuse is, unfortunately, one aspect I mainly leave aside for this chapter.

blocking online—that is, why counter-speech can't effectively do the blocking work it is supposed to do.

Beyond showing the limits of counter-speech in this case, this also demonstrate how the accommodation model relies, implicitly, on an overly rational and psychological picture in considering how content gets added to the common ground. But taking note of the red herring that psychology offers is a first step to seeing the problems with that account. This is not to say that psychology doesn't matter. My point is rather that when we take seriously the psychological state of those *targeted* by online abuse, we see that 'blocking speech' doesn't effectively block what it claims to block. And this is true of those doing the harassing as well, who often don't even acknowledge that any counter-speech even occurred, and instead carry on *as if* it hadn't, as is demonstrated in the conversational moves they continue to engage in.²⁰⁹

Again, this is not to say that counter speech is pointless or serves no purpose. It is simply to show its limits, and how those limits expose conceptual problems with the accommodation approach, especially with regards to online abuse. Moreover, this lays bare another common misconception shared between online and IRL subordinating speech, namely, that the right response is either the cliché about 'more speech,' or to simply 'ignore the trolls.' While these recommendations may seem at odds—one calls for a response while the other cautions against it—what they share is that each offers an individualistic solution that puts the onus on the victim of abuse.²¹⁰ More to the point, we can see how both are ultimately dismissive because both ignore the actual function of these speech acts. As Quinn puts it, "[s]ilence in the face of abuse is

²⁰⁹ For further evidence of this, I again refer the reader to any one of the number of first-personal accounts of living with online abuse, referenced above.

²¹⁰ In truth, the 'more speech' recommendation need not be individualistic, though it often is. See Gelber (2012) for a non-individualistic option.

not a solution; it's what abusers want" (2017, 50). Seen in this way, to ignore the trolls is to let the loudest voices win.²¹¹ But to treat this speech as if it were a more standard conversational contribution is just as likely to fail. Whether it is online or IRL, trying to respond with a reasoned argument is unlikely to change hearts and minds (Nielson 2012), and can in fact be quite dangerous (Delgado and Stefancic 2004).

So, considering the apparent inability of these existing accounts to easily explain the subordinating force of abusive speech performed online, I believe we need to look elsewhere. What's needed is an alternative account that can explain how seemingly non-authoritative and often anonymous speech can attain authority. As I argued in the last chapter, I believe a much more active process explains how some speech acts attain subordinating authority than the fairly passive model accommodation offers. These are cases I call *collective speech acts* as they are backed up by a collective authority attained by a chorus of speakers. Online abuse, I will argue, demonstrates this well.

4. ONLINE ABUSE AND THE CONSTRUCTION OF COLLECTIVE AUTHORITY

Online abuse can seem to capture a lot of different instances, and it's necessary to specify what's included under this concept. And intuitively, not every instance of online abusive speech, considered on its own, is of equal seriousness and harmfulness. Replying to a tweet with "Wrong!!! #MAGA" is one thing, and uttering death and rape threats is something else entirely. But there is a continuum and a pattern of escalation that connects the two.²¹² And these patterns

²¹¹ See Cherry (2015a) for a short argument as to why those harassed online ought not to go silent.

²¹² After distinguishing between the 'content' of online harassment and it's the 'behavior' it constitutes, Sarah Jeong (2018) gives a useful analysis of this spectrum, along with why we should think of it this way: "On one end of the behavioral spectrum of online harassment, we see a drive-by vitriolic message thrown out in the night; on the other

reveal illuminating facts about when and why a person's offline identity becomes relevant to their online persona. Like other forms of harassment, online abuse tends to pick out and objectify an individual based on markers of identity, and thus enables the coercive reduction of a person to (only) some aspect of their public presentation. It can be somewhat striking, then, in what ways an individual's offline identity matters, or doesn't, to the harassment they receive.

There is the plain fact that if one is a member of an oppressed group offline, then that matters for how likely they will suffer abuse online, and, of course, what form that abuse will take. As research from the Women's Media Center Speech Project confirms, women are more likely to be victims of online abuse, and the content of that abuse is overtly misogynistic.²¹³ Men and women of color often receive racist comments in response to mundane posts, especially if they are public figures. This, itself, is not that surprising. Those who are already targets of oppression find themselves targeted even more. And, as Quinn writes about her experience running a crisis line for victims of online abuse: "For all the promises of meritocracy and the internet allowing minds to transcend our floppy meat bodies, it's exceedingly rare to see a case of online abuse that doesn't tie back to the client's offline body" (2017, 203). That is to say, the advent of the digital age has not led to the irrelevance of identity, as some early advocates predicted.

end, we see the leaking of Social Security numbers, the publication of private photographs, the sending of SWAT teams to physical addresses, and physical assault. Saying that these behaviors are all on the same spectrum does not mean that they merit the same kind of censure and punishment. [...] I don't think that placing these behaviors side by side means it's impossible to differentiate between one act and the other for the purposes of post hoc punishment. Seeing these behaviors on the same spectrum becomes illuminating because it teaches us not how to punish, but how to design environments to make targets feel safe."

²¹³ See the Women's Media Center's Speech Project report for the data, including "an 11-year analysis of online harassment cases found that women made up 72% of victims and men 47.5% of perpetrators." And see Manne (2017) for an account of misogyny that takes seriously this victim-centered form of online abuse.

In these cases, a person’s identity is what marks them for abuse, at least in part, and no doubt for what kind of abuse they receive. Something different, though related, occurs in cases of doxing (publishing private and identifying information online without consent) and swatting (calling in fake reports of a crime to the police with the intention that they send a SWAT team to a target’s address)—and dead-naming in the case of trans targets of abuse—where harassers use a target’s offline life as a weapon against them.²¹⁴ The details of their offline identity are stolen from them. Private facts quickly become much too public.²¹⁵ So here we see two roles a target’s identity can play in online abuse: first, as a factor that either motivates the harassment or informs the type of harassment; second, as something to be exploited by harassers.

Yet this pairs somewhat uneasily with another common feature of online abuse, a feature it shares with another old act with an interesting rebirth in the digital age—online shaming. In both online shaming and online abuse (two distinct acts with a clear overlap), there is an implied need to build up target’s status in order to justify the importance of taking them down. Many victims of online abuse, and especially in cases that ‘go viral’ and bring in anonymous abusers, report similar false claims about their social status—like being secretly rich (Quinn 2017; Ronson 2015). In practice, claims like these are taken as reason enough for the toxicity now coming their way. As Kathryn Norlock (2017b) explains: “The target of the shame is the one that needs to be ‘taken down,’ a phrase that implicitly indicates the comforting narrative that the

²¹⁴ It may seem odd to include doxing—disclosing personal information online against a person’s will—here. However, as Sarah Jeong (2018) suggests, “like online verbal abuse, doxing is a tactic to dominate the voice of the internet,” and so fits in alongside other forms of online abuse.

²¹⁵ Privacy is the focus Levmore and Nussbaum’s edited collection, *The Offensive Internet*, in the introduction of which, they note: “a spotlight on privacy clarifies the novelty of the Internet. A bit of information once thought confidential may now blanket the globe with the help of the Internet; a false and defamatory accusation about a person may become a constituent part of that person’s Internet identity, where it affects relationships and employment opportunities for all time.” (2010, 9). For the threat to privacy represented not by abusers but by government and big business, see Schneier (2015).

target of shame is up high, in a position of power that deserves to be brought down a peg.” And an important corollary of this is, as Norlock says, that “the shamers see themselves as the disempowered.”

What I want to suggest unites much of the above—targeting members of oppressed groups, making them feel unwelcome in online communities, and the invasion of privacy into someone’s “real” life on the one hand, along with the complete disregard for what that “real” life actually entails on the other—is the construction of ingroup and outgroup status. In these cases, online abusers are targeting someone who is, in brief, *not like them*—or at least constructed as such.

And, of course, this is all the more frightening as the perpetrators of this sort of online abuse are often anonymous. As we saw above, though, anonymity in online abuse represented one point of departure from IRL hate speech and posed problems for existing accounts of the force of such speech. In the next section I argue that anonymity is used by harassers to construct the image of group solidarity in the target’s mind.

4.1. Anonymity and the Force of (Veiled) Threats

As everyone knows, threats of physical and sexual violence are not rare online. Indeed, one of the common tropes of e-bile that Jane highlights is the ubiquity of violent misogyny:

E-bile targeting women commonly includes charges of unintelligence, hysteria, and ugliness; these are then combined with threats and/or fantasies of violent sex acts which are often framed as “correctives.” Constructions along the lines of “what you need is a good [insert graphic sexual act] to put you right” appear with such astounding regularity, they constitute an e-bile meme. Female targets are dismissed as both unacceptably unattractive man haters and hypersexual sluts who are inviting sexual attention or sexual attacks. (Jane 2014, 533)

And while direct threats do occur, more common is violent aggression expressed in the form of “hostile wishful thinking, such as ‘I hope you get raped with a chainsaw’” (Jane 2014, 533).²¹⁶ While this indirect phrasing allows abusers to avoid legal trouble and skirt Terms of Service, it doesn’t make these statements any less threatening to their targets. If anything, it is often an escalation, as it seems to imply a coordinated group-effort with a division of labor.

That is, veiled threats of this sort are only properly understood when we consider them in their fuller context, where they tend to imply a larger network of abusers. First, if the threat comes from an anonymous or unknown account—a non-follower, for instance—that often implies they were likely directed there by others, as the coordination of abusive campaigns with a sort of division of labor is not uncommon. As Sarah Jeong reports,

[the] examination of sustained harassment campaigns shows that they are often coordinated out of another online space. In some subcultures these are known as ‘forum raids,’ and are often banned in even the most permissive spaces because of their toxic nature. In the case of the harassment of Zoë Quinn, Quinn documented extensive coordination from IRC chat rooms, replete with participation from her ex-boyfriend. (2018)²¹⁷

Second, given that this abuse commonly occurs on a public platform, where it is viewable by anyone, the way these utterances are given uptake reveals something important about the speaker and the wider community. As Lynne Tirrell argues “our speech acts also undertake a meta-level expressive commitment about the very saying of what is said. Expressive commitments are commitments to the viability and value of particular ways of talking (Tirrell

²¹⁶ Sarah Jeong (2018) calls this “Colorably Threatening Harassment”, which is: “Harassment that is not overtly threatening, but is either ambiguously threatening such that an objective observer might have a hard time deciding, or is clearly intended to make the target fearful while maintaining plausible deniability.”

²¹⁷ While this is only one instance, further evidence suggests this practice is not as uncommon as some presume. Tufekci (2017) discusses a coordinated harassment campaign she discovered against her, and Gray-Donald (2018) discussed the case of Nora Loreto. As Whitney Phillips notes however, “Harassment campaigns and other coordinated attacks are particularly challenging to report on responsibly, as the entire purpose of these attacks is to generate oxygen, draw more people into a story, and create the biggest public spectacle possible” (Phillips 2018).

2017, 144). When these expressive commitments aren't uniformly condemned, they can shift the boundaries of what counts as acceptable discourse in a community, opening the door to further harm (Tirrell 2017; Saul 2017). I discussed one aspect of this in the previous chapter as I made the initial case for the more active role an audience can play in *collective speech acts*. In the case of online abuse, given that harassing speech in this medium often receives 'likes' from other users, these features—commitments to the value of this discourse and shifting boundaries of permissibility—take *tangible* form.

What might at first glance seem like a one-off message from a single individual can, in fact, reveal a group of like-minded people whose 'real' identities, and thus their relationship or proximity to the target, are hidden from view. It is in this sense that it is a mistake to view the speech acts typical of online abuse in an individualistic lens, separate from their wider context. This is not too difficult either; it simply amounts to listening to those who have experienced this harm. As Jeong (2018) says, "targets of harassment, particularly members of marginalized groups, may view a single comment differently than an outsider might, because they recognize it as part of a larger pattern."

For those targeted by such speech, then, it can understandably be read as a glimpse into the ingroup speech of others, where marching orders are being given, well-received, and which might then be carried out by any one of the many anonymous figures on the other end of internet.²¹⁸ This takes a very real toll on its targets. As Lindy West writes about her experience:

online harassment is not virtual—it is physical. Flooding in through every possible channel, it moves and changes my body: It puts me on the phone with the FBI, it gives me tension headaches and anxiety attacks; it alters my day-to-day behavior (Am I safe? Is

²¹⁸ The increased attention given to 'incel' violence against women shows how this notion of 'marching orders' is not too far from reality. In these cases, it's not unfair to say that some of these killers were persuaded to murder within online communities—indeed this is the meaning of the 'incel' term "going blackpill." For more, see ContraPoints (2018); Beauchamp (2019).

that guy staring at me? Is he a troll?); it alienates my friends; it steals time from my family. The goal is to traumatize me, erode my mental health, force me to quit my job. (West 2016)

As this makes clear, while anonymity poses challenges for the description of online abuse—namely, by foreclosing the standard explanations for the authoritative force of subordinating speech—it in fact provides a powerful tool for those who wish to inflict harm on their targets.²¹⁹

It is the combination of anonymity and apparent group solidarity—‘likes’ instead of condemnation—that is a dangerous mix for targets of abuse, and I claim an important source of the authority these speech acts rely on to subordinate. This is evident in Quinn’s description of her own experience with online abuse, “I read many of the threats in my ex’s voice,” she says. “But this was somehow more insidious—he wasn’t just continuing his abuse; *he was crowdsourcing it* (2017, 51).²²⁰ This vivid account is supported by media researcher Eden Litt, who notes that “without being able to know the actual audience, social media users create and attend to an imagined audience for their everyday interactions” (2012, 333). That is to say, when we can’t directly perceive our audience, we create it in our minds.

Norlock adapts Litt’s insight to the case of online shaming. “Recipients of shaming, like Justine Sacco,” she says:

are also living in the web of imaginal relationships. And imaginal relationship literature suggests that her shamers also *live in her head*, as she too may be forced to rely on her imagination and mentally conceptualize her detractors. In such situations, advising a victim of shaming to “ignore the trolls” is beyond pointless. [...] The advice to ignore the social community as it lives in one’s head is more than ineffective—*it’s missing the force*. We are unable to control the extent to which intangible words in cyberspace take

²¹⁹ Ronson suggests this is the case for online shaming as well: “‘Social media shamings are worse than your shamings,’ I suddenly said to Ted Poe [a former Judge who often used public shaming as part of his sentencing]. He looked taken aback. ‘They are worse,’ he replied. ‘They’re anonymous.’ ‘Or even if they’re not anonymous, it’s such a pile-on they may as well be,’ I said” (Ronson 2015, 88).

²²⁰ Though, it’s worth noting that neither Quinn or myself is claiming that anonymity is the *source* of harmful speech and therefore ought to be prohibited in all online spaces. As Quinn says: “Online anonymity isn’t responsible for the prevalence of horrible behavior online. Shitty moderation is” (quoted in Jeong 2018). For more on anonymity’s role in online speech, see Belmont (2018).

the form of imaginal relationships that burden or brighten our self-perceptions. (Norlock 2017b, emphasis added)

Even though anonymity forecloses some standard accounts of speaker authority—like those afforded by race and gender—by seeing how anonymous avatars can become a monolith in one’s mind we can add a bit more detail to a conception of group-authority that in fact requires something like anonymity. In leveraging the target’s own cognitive resources—namely, their capacity for imaginal relationships—large-scale online abuse campaigns become more than the sum of their parts. Anonymity creates the semblance of cohesion where there might not in fact be any, uniting different speakers who might not have anything in common or prior relationship aside from their hostile speech directed at the same individual.²²¹

Furthermore, it is *through* this speech that they become united (at least in the mind of the target). I therefore claim this is a case of *collective speech*, whose authority is sustained by the active participation of a community. Recall from Chapter Three:

Collective authority: Through repetition and endorsement, individually non-authoritative speech is made authoritative by signaling collective support and solidarity. A group of speakers have their authority secured through a repetitive process where each new utterance adds to the strength of the overall practice.

In online abuse, ‘likes,’ ‘retweets,’ hashtags, and more play the role of active community endorsement, adding some authority to speech that otherwise lacks it. Like accommodation, audience uptake secures authority for speech that, absent that uptake, would have a different pragmatic force. But as I have emphasized, in these cases, the practices that do the heavy lifting here are *active*, not passive. This makes sense, moreover, as the process I am attending to is not

²²¹ It is also revealing how, in some cases, when anonymity is removed, so is some of the speaker’s power. The story of the anonymous poster who created the ‘Trump Slams CNN’ gif, and his subsequent apology after his ‘real’ identity is uncovered one interesting example of this dynamic. For context, see North (2017).

primarily psychological, but social. It is the active construction of ingroups and outgroups. And while these can occur passively—for example, when it slowly dawns on someone that they are now in a committed relationship—more common are active process of initiation. And even in the more passive cases, as Herbert and Kukla point out, there is nonetheless an active, social component that occurs when one is recognized as an insider. As they say:

being recognized as an insider by insiders is not just the recognition of a separate fact; rather, this recognition plays a constitutive role in having that insider status. Part of being an insider is being recognized as one. Crucially, the relevant sort of recognition is not mere passive, conscious acknowledgment, but the kind of recognition that is built into practice. (Herbert and Kukla 2016, 584)²²²

In online abuse, this takes the form of harassers cheering on other harassers, directing them to new targets, sharing information, one-upping one another; etc. In all these cases, speech plays an active role in solidifying the collective authority that strengthens their words, turning it into the genuinely subordinating speech that it is.

4.2. Shared Language and Solidarity

Above, I argued that in cases of online abuse, anonymous speech is not, in fact, identity-free speech. Indeed, anonymity can contribute to the active construction of a group-identity that may be wielded to inflict great harm. But anonymity is only part of the explanation; shared, insider language is the other. In this section I strengthen the argument that shared language plays a vital role in constituting the type of collective authority at issue in online abuse.

²²² For more, see Herbert and Kukla, who argue that “ingrouping, outgrouping, and displaying, negotiating, and constituting community boundaries are among the important pragmatic functions of language” (Herbert and Kukla 2016, 579).

To start, it's useful to note that the affordances of social media make it clear how a person's speech is always tied to their (ever shifting) socially constituted position—even when it's anonymous. Whether via a profile picture, a short bio, a hashtag, or an emoji, social media brings new means of signaling or reading one's identity, along with one's membership or allegiance to broader social groups. I want to emphasize however, how this just amplifies what is also the case offline. My favorite articulation of this comes from Mary Louise Pratt, who wrote:

once you set aside the notion of speech acts as normally anchored in a unified, essential subject, it becomes apparent that people always speak from and in a socially constituted position, a position that is, moreover, constantly shifting, and defined in a speech situation by the intersection of many different forces. On this view, speaking “for oneself,” “from the heart” names only one position among the many from which a person might speak in the course of her everyday life. At other points, that person will be speaking, for instance, as a member of some collective, or as a rank in a hierarchy, and so forth. Nor is there any guarantee that these positions will be internally consistent or consistent with each other. Everyone has had the experience of being situated among complex and contradictory forces, as, for example, press secretaries stationed between the revelation demanded by their addressees and the concealment demanded by those for whom they speak, or the underling required on the one hand never to correct superiors so that they will not look stupid, and required on the other hand always to correct superiors so that they will not look stupid. (1986, 63)

On social media, these implicit features made fully explicit, often purposely so. What's true of offline speech becomes heightened in online speech in illuminating ways. A profile picture can reveal aspects of a person's identity that alter how their speech is interpreted, along with what speech act they are performing. Including a Pepe emoji or #MAGA, for example, can instantly communicate one's broader allegiances and provide all the necessary context an audience needs to decipher something important about the speaker and their utterance, like whether it was intended as constructive criticism or an insulting dismissal. These aspects of online speech can be like wearing a pin or a badge and allows individuals to actively construct and manage the version of themselves they present to (one segment of) the world. This allows for a lot of variety, freedom, and play, including the inconsistencies that Pratt describes—e.g., concealing or

emphasizing distinct parts of oneself for different platforms, speaking *as an x* on one account and *as a y* in another, going fully anonymous for personal safety, or using ‘alts’ for whatever reason.

All this is simply to say that a major component of the discourse on social media is structured around actively constructing a persona that is (a version of) ‘you’ and interacting with other, similarly self-constructed others. And as Michael Lynch points out, if we conceive of the self as constituted through narrative construction, this suggests, then, that “we are stories that are increasingly constructed online in social networks” (2016, 74). Thus, the management of one’s own identity, along with the interpretation of others, sits at the foundation of nearly all online discourse, fixing a lot of facts about how a discourse will unfold before any claims are even uttered.

What’s relevant for my purposes is how, on social media, this type of signaling often occurs by citing the speech acts of another. “Speaking for oneself,” in this context, often means speaking with the voice of another. This is in part how one builds their chosen identity. While this is not an uncommon feature of speech, it is heightened and made explicit online. For instance, in its early days, Facebook used to ask for your favorite quotes that it would then display prominently in your profile. While that specific practice is dated, similar tropes are still common. Large and powerful social movements can galvanize around a hashtag that, in essence, consists in joining with the voices of others. As examples like #BlackLivesMatter, #NeverAgain, #MeToo, #ICantBreathe, and even #MAGA demonstrate, a chorus of online (and offline) speech acts can be a rallying call that affects real change—though, not always the specific change we would hope for.²²³ At societal level, this can bring out both good avenues for effective

²²³ See Tufekci (2017) for an analysis of the impact of social media and other digital communication technologies on progressive activism, as well as how repressive regimes have learned to clamp down on these groups. And see

solidarity—where people are able to amplify and signal-boost others in productive ways—and bad ones, as practices like the co-opting and appropriation of the words and voices of the more marginalized are all too common.²²⁴ And at the interpersonal level, this is relevant for the case of online abuse because, as Quinn (2017) puts it, “the same techniques that people have used to organize important grassroots movements like Black Lives Matter can be used by people trying to destroy someone” (52).

Here, I want to focus on one technique that is crucial to the development of the group authority at issue in online abuse. Namely, the use of hashtags and other similar rhetorical constructions to unify the voices of many into an ad hoc collective. As I will describe it, hashtags are *explicit ventriloquisms*, and so are a vivid example of language’s role in constituting a group identity.

In the course of building his account of slurs, the linguist Geoffrey Nunberg (2018) describes *ventriloquisms* with the following:

In a particular context, a speaker pointedly disregards the lexical convention of the group whose norms prescribe the default way of referring to *A* and refers to *A* instead via the distinct convention of another group that is known to have distinct and heterodox attitudes about *A*, so as to signal his affiliation with the group and its point of view.

That is, when a speaker uses a *ventriloquism*, they are pointedly disregarding the standard term or terms that the convention dictates and are instead pointedly using the voice of another. In doing so, they signal their allegiance to a specific community, at least in that moment. Nunberg

Khan-Cullors and bandele (2018) for the story of how social media played a key role in the growth of Black Lives Matter.

²²⁴ The phrase and hashtag ‘Black Lives Matter’ has been taken up, twisted, and put to use for all sorts of ends, including opposing forces (e.g., #AllLivesMatter, #BlueLivesMatter) and clueless vegans (e.g., #AnimalLivesMatter, #ElephantLivesMatter). I discuss some of these dynamics in the next chapter.

uses the example of a university dean using *ain't* in place of *isn't* to implicate that the knowledge being communicated was more folksy than academic.²²⁵

While Nunberg's main goal is to argue that slurs are cases of ventriloquisms, I don't want to commit to anything like that. But I do want to suggest that his analysis clarifies the pragmatic force of hashtags (and other uses of identical phrasing) in online abuse. As I will put it, hashtags function as explicit ventriloquisms, and in doing so serve to strengthen shared group identity for harassers. And as I suggested above, this group identity is a crucial element of online abuse, as it aids in the construction of the collective authority at issue in online abuse.

What I am pointing to is part of a larger phenomenon where non-slurring terms are thought to mimic to some degree how slurs are deployed, especially in online discourse. This is what leads Jonathan Chait to claim ludicrously that 'neoliberal' is a slur.²²⁶ To me, that is a bit of a joke. But the function of terms like 'SJW,' 'feminazi,' 'cuck,' 'globalist,' 'lib,' 'elite,' 'TERF,' 'deplorable,' 'lanyard,' and more demonstrate some interesting flexibility in language, and in particular how quickly a rigid sentiment can galvanize around a term online—especially ones used to label outsiders. Whatever you may think about these cases—and, to be clear, I don't think any amount to being slurs—my basic point is that there is something noteworthy occurring in the use of things like hashtags in online discourse, and one way to examine this is via the analysis of slurs in terms of ventriloquisms that Nunberg puts forward.

According to this account, slurs have the force they do not because of any semantic properties, but because of what the choice to use them says about their speakers. As he puts it,

²²⁵ As Nunberg explains the case of a “dean at an Eastern university [who said]: ‘Any junior scholar who stresses teaching at the expense of research ain't gonna get tenure.’ In the dean's mouth, the use of the demotic *ain't* rather than *isn't* implied that his conclusion wasn't based on expert knowledge or a research survey; it was as if to say, “You don't need an advanced degree to see that; it's obvious to anyone with an ounce of sense.” (2018)

²²⁶ To be precise, he claimed it is a “term of abuse.” See Chait (2017). Or more recently Howard Schultz suggested the same of the term 'billionaire,' saying that 'people of wealth' is a better, less divisive term.

“To use a slur is [...] to signal one’s affiliation with a group that has a disparaging attitude towards the slur’s referent” (2018). “In a nutshell,” he says, “racists don’t use slurs because they’re derogative; slurs are derogative because they’re the words that racists use” (Nunberg 2018).

That is to say, the crucial piece of information that a speaker communicates when they use a slur, according to Nunberg, is that they willingly wish to be associated with a group of other’s speakers whose attitudes about the slur’s target they share. In other words, slurs say more about the group of people who use them than the group to whom they refer. Crucially it is not just shared attitudes that are implicated, but shared group membership.²²⁷ This matters for an account of slurs according to Nunberg because:

As [Langston] Hughes tells it, the force of *n****r* goes beyond anything the speaker believes or feels about blacks, or for that matter, beyond anything that others who have used the words have thought or felt about blacks. It also evokes the things such people have *done* to blacks—with the speaker pointedly affiliating himself with the perpetrators. *The word can turn a bigot from a hapless, inconsequential 'I' into an intimidating, menacing 'we.'* (Nunberg 2018, emphasis added)

While I am not willing to commit to this account of slurs, I do see something importantly correct here in the analysis of subordinating speech acts. Namely, the construction of a collective identity through shared language and tropes. Renée Bolinger (2017) develops a similar pragmatic account of slurs that can add to this story.²²⁸ In particular, Bolinger explains the (varying) offensiveness of slurs via a *contrastive choice* account. This adds to the idea of ventriloquism by further explaining how marked expressions can carry important signals about their speakers.

²²⁷ This, according to Nunberg, distinguishes his view from a similar one offered by Elisabeth Camp. See (Camp 2013).

²²⁸ I see it as similar to Nunberg’s in part because it is a pragmatic account where the semantics do little in explaining the relevant features of slurs.

As Bolinger puts it: “when we use slurs, we communicate information about ourselves and our attitudes towards the targets” (2017, 439). This information is signaled, moreover, through a speaker’s decision to a particular term over a non-marked alternative. As she explains:

For signals based in *contrastive choice*, the relevant behavior is the free selection of a marked expression, and performance signals that the speaker endorses a cluster of attitudes associated with the term (or, more precisely, a high probability that the speaker shares some or all of the attitudes in this cluster). (Bolinger 2017, 447)

Consider the use of the term ‘SJW,’ particularly as it occurs online. This is, in most cases, used pejoratively, referring commonly to individuals who promote socially progressive views like feminism, anti-racism, etc. Importantly, this term is used almost exclusively by those who *oppose* these goals, or at least as they promoted by the SJWs. In using this term then, whether prefixed by a hashtag or not, speakers pragmatically convey information about their own group-membership. Again, as Bolinger explains:

The information content of signals based in contrastive choice is linked to how marked the term is: if α is a term that is used almost exclusively by speakers who embrace φ , and this fact is well-known, then a contrastive preference for α is a high-information signal, raising the probability of the speaker’s endorsing φ nearly to 1. The more well-known the association between α and φ is, the higher the information content of the signal, and thus the more strongly the contrastive choice signals the speaker’s endorsement of φ . (2017, 447)²²⁹

Since ‘SWJ’ is used mainly by its detractors, and since this is well-known, using it carries a high-probability signal that the speaker endorses the views too. That is, what terms like this do, is to express and solidify group membership. This is a dynamic process performed primarily through speech acts. And it is through this process, moreover, that the targets of online abuse come to recognize that they are being addressed not by a single speaker, but by a mob. This

²²⁹ Moreover, on Bolinger’s view, this is not reducible to speaker intentions: “Signaling on this framework is factive: a speaker signals some content φ when her use of an expression satisfies the conditions, regardless of whether she intended to communicate φ , and independent of whether hearer uptake occurs” (Bolinger 2017, 447).

interpretation makes sense, since it is often exactly what is occurring. And as Jeong reports, it is this interpretation that makes of the “really bizarre phenomenon” of “all the low-level mobbers, who have little-to-no real investment in going after the target, and would not manifest any obsessions with that particular target without the orchestrator to set them off.” As she explains:

Here they resemble the zombie nodes of spam botnets, right down to the tactics that have been observed to be deployed—rote lines and messages are sometimes made available through Pastebin, a text-sharing website, and low-level mobbers are encouraged to find people to message and then copy/paste that message.

Here again we see how in online abuse, the implicit is made fully explicit. Speakers are literally copying and pasting their utterances from one another, and in doing so adding strength. More importantly, this shows vividly why an individualist approach to online abuse is inapt for describing the force of these speech acts.

It is only when we see these speakers as part of a collective, and a collective, moreover, that is constructed in part through the active use of shared speech acts, that we capture the pragmatic impact of these speech acts. They are, in an important sense, *collective speech acts*, and it is in this way that they achieve their primary function of subordinating their targets.

5. CONCLUSION

In this chapter I have considered some of the key features of online abusive speech. I argued that anonymity plays a key role in building a type of collective authority for online abusive speech acts, and moreover, that the endorsement of a group identity through shared rhetorical constructs like hashtags further adds to the targets’ sense that they are being addressed by a collective rather than individuals. I argued that this combination of anonymity and apparent group solidarity—shared phrases and hashtags, ‘likes’ and retweets—is a dangerous mix for

targets of abuse, and I claim an important source of the authority these speech acts rely on to subordinate. That is, it is in this way that online abuse becomes more than the sum of its parts. These and other features build collective authority for seemingly isolated speech acts and, as I will now suggest, reveal aspects of subordinating speech that are often overlooked.

In her analysis of online shaming, Kate Norlock suggests that in addition to the unique problems brought on by anonymity, the “speed and volume of online affirmation outmatches what the human mind evolved to manage.” From this she argues that our online interactions give rise to important new responsibilities (2017b). While I agree that some features of most online abuse make it unique to an extent, I believe that greater attention to similar features could help to highlight under-emphasized aspects of IRL hate speech. I’ve emphasized the role of anonymity in cultivating the appearance of coordination and a division of labor in online abuse—even if they’re in fact is none—but the same can be said about IRL hate speech. Racist graffiti spray-painted on college campuses, slurs yelled from passing cars, white-nationalist flyers and rallies proudly displayed in public all invoke anonymity and group-authority in a similar way to create an overall environment of exclusion. It is for this reason that I think that collective speech acts are a broader phenomenon worthy of investigation across mediums. Accounting for the realities of subordinating speech both online and IRL shows that speaker authority is dynamic and emergent, and often depends on the wider community in more ways than simple accommodation. Finally, seeing hate speech for the group activity it is reveals the limitations of the ‘more speech’ / ‘don’t feed the trolls’ paradigm in both online and IRL contexts. Focusing on individual instances isn’t enough, and rejecting this individualistic lens is also necessary for devising solutions to the harm they present. As Max Fisher and Amanda Taub report:

It is becoming increasingly common for groups of people, whipped into a rage by influential people on social media, to single out targets for mass campaigns of online

harassment and threats. [...] The main problem seems to be that social media companies' guidelines tend to focus on content in isolation. Because the accounts that instigate the hatred and rage don't necessarily participate in the mass harassment directly—often their followers are the ones who send the death threats or do the doxxing—this problem is a poor fit for that approach.” (Fisher and Taub 2019)

As this shows, tackling this problem properly requires addressing the collective from which the speech draws its power, whether it is an organic ad hoc group, or, more commonly, a pre-existing community with a clear (if informal) hierarchy. Seeing this bigger picture is helpful in explaining the damage it can do to a community. Online abuse reveals this group-based authority well, but this is a feature shared by IRL forms of subordinating speech, and one that must be kept in mind. And yet, given that the dominant mode of understanding subordinating speech both online and IRL is individualistic, this is too often obscured from view.

CHAPTER FIVE: THE PRAGMATICS OF PROTEST AND POSITIVE PROPAGANDA

“I think it’s silly when people want to add things to [the phrase ‘Black Lives Matter’] or feel worried that it’s too divisive. What’s divisive is cops killing black people. What’s divisive is vigilantes shooting up churches. That’s divisive.”

- Patrisse Khan-Cullors, co-founder of Black Lives Matter²³⁰

1. INTRODUCTION

In this final chapter, I take up another instance of what I’m calling *collective speech acts*: the speech acts of protest. But rather than focus on the hateful protests of bigots like the Westboro Baptist Church, I will turn to specifically egalitarian, liberatory protest, using Black Lives Matter protests as my main example. I do so, in order to show how this is also a case we should understand as a collection of disparate utterances combining to produce a distinctive type of group-based speaker-authority, opening powerful avenues for collective action—as the history of social movements show.

Without question, the phrase ‘Black Lives Matter’ has cemented itself as a political speech act of great importance. Media discussions on Black Lives Matter (BLM) protests often focus narrowly on the slogan that gives the group their name. It therefore offers a unique route to a greater understanding of the pragmatics of protest as a speech act.

By reflecting on the significance of BLM, I argue that it is the distinct pragmatic features of protest—its entitlement conditions, and the uptake it aims at—that best reveals its moral, political, and epistemic significance. In short, I claim that, much like some forms of

²³⁰ See Bell and Kondabolu (2017).

subordinating speech, protest is a clear example of socially located, collective speech. And this is because an essential aspect of protest is its function of foregrounding the moral authority of the speakers. The conception authority here, moreover, must account for the basic moral status of the speakers, their situated knowledge, as well as the contextual and interpersonal relations that inform the “total speech situation,” as Austin (1962) would say. These elements shape the pragmatic force of protest.

To better approach these issues, I consider Jason Stanley’s account of “positive propaganda,” which he takes to be an important avenue of anti-oppressive resistance. I evaluate his suggestion that some instances of protest—such as the 1964 march on Montgomery—are paradigmatic examples of positive propaganda (2015, 113). I argue that Stanley’s model is an unhelpful tool to apply to most forms of protest, and an examination of BLM protests illustrates why. This is because, as I argue, we lose sight of what makes protest distinct if we think of it as engaging in argument. But, a focus on content, even unarticulated content, leads us to evaluate the claims of protesters separately from the context that produced them. Stanley’s model encourages this abstraction, as does the greater attention it gives to the audience in place of the speakers. On his account, the crucial role of the speakers is obscured, as is the relationship between speakers and audience. In this way, by using Stanley’s account of propaganda as a starting point, I consider what protest is from a pragmatic point of view, and how it relates to propaganda and subordinating speech more generally.

One guiding idea for this chapter is that we need to examine the interaction between the protests of social movements on the one hand, and counter-protest speech that arises in response on the other. I’m concerned with how these two phenomena occur in tandem and want to show how we get a better understanding of each type of act by looking at them together, rather than in

isolation. I argue that seeing each of these acts as collective speech acts, and furthermore as interacting with each other, helps to reveal their differences more effectively than does analyzing each independently of the other. I locate the differences between these speech acts not only in their content, nor the ends they seek, but in their whole social situatedness. My goal, therefore, is to distinguish the righteous moral protest of the indignant from the dismissive and reactionary counter-protest speech it inevitably generates. And as I argue, it is by seeing these acts in their proper context that we're led to appreciate the details and dynamics of both subordinating speech and liberatory protest.

2. PROPAGANDA: POSITIVE AND NEGATIVE

In *How Propaganda Works*, Jason Stanley examines a sometimes subtle, but nonetheless dangerous type of subordinating speech. According to Stanley, propaganda is dangerous because of how it can erode important political values in liberal democracies (Stanley 2015, especially Chapter Three). To get a firmer grasp on this claim, let's consider some useful distinctions and examples.

Stanley discusses propaganda as it is a “contribution to public discourse that is presented as an embodiment of certain ideals, yet is of a kind that tends to erode those very ideals” (ibid., 53). Climate change denial campaigns provide a useful example (ibid., 60). When oil companies promote the views of the few climate scientists who deny anthropogenic climate change, they engage in this type of *undermining* propaganda, because they appeal to a worthy ideal—scientific objectivity—in the service of a goal that tends to undermine that ideal.

An illustrative example of this is offered by the author Michael Crichton (of all people) who apparently once said in a speech:

Let's be clear: the work of science has nothing whatever to do with consensus. Consensus is the business of politics. Science, on the contrary, requires only one investigator who happens to be right, which means that he or she has results that are verifiable by reference to the real world. In science consensus is irrelevant. [...] The greatest scientists in history are great precisely because they broke with the consensus. There is no such thing as consensus science. If it's consensus, it isn't science. If it's science, it isn't consensus. Period. (Crichton 2003)

How a statement like this, which on its surface seems to embody the ideal of scientific objectivity, can actually serve to *undermines* this ideal, is clear. The mistrust it aims to generate in scientific practice erodes trust in science, and fuels an unhealthy skepticism about climate change research.

When the ideals are specifically *political* values, Stanley calls it “demagoguery,” which is:

A contribution to public discourse that is presented as an embodiment of a worthy political, economic, or rational ideal, but is in the service of a goal that tends to undermine that very ideal. (Stanley 2015, 69)

Demagoguery is always antidemocratic because “it wears down the possibility of democratic deliberation” (ibid., 82). Arguments for racist voter ID laws that exploit “ideals like ‘one man, one vote,’ together with the appeal to voter fraud,” are a clear example of this type demagoguery, as well as its undemocratic effects (ibid., 68–69).

One political ideal that Stanley discusses in detail is the ideal of *reasonableness* in the realm of public reason—that is, the norms and standards that ought to guide public discourse. Following the ideal of reasonableness here means that debate and discussion about matters of public interest are “guided by equal respect for the perspective of everyone subject to the policy under debate” (ibid., 94). At the heart of this norm of reasonableness is the capacity for

empathy—the ability to imaginatively position oneself in the situation of another. According to Stanley, to say that public discourse is guided by reasonableness is to say that public policy discussions don't exclude the voices anyone who might be affected.

On the assumption that something close to reasonableness is an operative ideal in a democracy, cases of propaganda will then typically look to make the appearance of being reasonable, while actually serving to make debate on a topic biased and unjust. It will do so by making it harder for some members of the community to participate on fair terms, in part by diminishing the capacity for empathy within dominant members of that community. As Stanley puts it, “paradigm cases of propaganda will be ones that represent it to be reasonable not to take certain perspectives into account” (ibid., 108). This can occur, he suggests, when propaganda presents “the perspectives of some of our fellow citizens as unworthy of consideration” (ibid., 122).

One method that makes this erosion of empathy possible is the prevalence of persistent negative stereotypes about certain social groups. Negative stereotypes of Black Americans, abetted by the use of terms like “superpredator,” can have the effect of eroding empathy and set the stage for anti-Black policies (ibid., 123). And this can occur alongside or even as part of seemingly reasonable contributions to a debate, in part by the foregrounding of ‘reasonable’ topics like public safety.

Here is a good place to note that while Stanley defines propaganda in terms of its *effects*, the *mechanism* of propaganda that he elaborates most clearly focusses explicitly on the expressed content of utterances. That is, according to the main mechanisms that Stanley's account describes, propaganda works by communicating propositions that strengthen or weaken certain

ideologies.²³¹ On Stanley’s model, utterances contribute new propositional content to the shared conversational background—the “common ground”—that then inform future linguistic and non-linguistic moves.

Crucially, utterances can contribute new propositions to the common ground covertly and indirectly. This often occurs through the careful use of presupposed or “not-at-issue content,” which enables the addition of new material to the conversational participants’ common ground without their conscious assessment of this material. For example, when a speaker says: ‘I spent part of every summer until I was ten with my grandmother, who lived in a working-class suburb of Boston,’ the location of the grandmother’s home is the not-at-issue content, while the summering habits of the speaker are the at-issue content. We see this by considering what someone would typically be denying if they challenged the speaker’s statement (2015, 134).²³² In this example, it would be where the speaker spent her summer, rather than whether her grandmother lived in a working-class neighborhood.

As Stanley says, the not-at-issue content of an utterance “is not advanced as a proposal of a content to be added to the common ground. Not-at-issue content is directly added to the common ground” (2015, 135). For example, Stanley argues that repeated associations of the term “welfare” with images depicting Black Americans as lazy has made the generic claim “Blacks are lazy” part of the not-at-issue content of the term “welfare” (ibid., 138). This means that a contribution to political discourse may be presented as an embodiment of reasonableness, but due to the use of the term “welfare,” it will also communicate not-at-issue content that erodes

²³¹ These mechanisms are discussed more fully in Chapter Four of Stanley (2015), while the general definition of propaganda appears in Chapter Two. It is worth noting however, as several reviewers have, that Stanley’s stated definition of propaganda seems to be inconsistent with his general use of the term throughout the book. See Wolff (2016); or Brennan (2017).

²³² Stanley takes this example from Potts (2005). For a criticism of the implications Stanley draws from this to the case of propaganda, see McKinnon (forthcoming).

reasonableness, further reducing empathy for Black Americans. This is what makes the skillful deployment of not-at-issue content fertile ground for propagandists and explaining this process in detail is Stanley’s main project in the book.²³³

2.1. Positive Propaganda

Let’s turn now to what Stanley calls “positive propaganda.” These, on the other hand, are contributions that *strengthen* democratic ideals. Following W.E.B. Du Bois (1926), Stanley calls this “civic rhetoric”, and describes this type of political speech as:

A contribution to a debate [that] can improve the subsequent reasonableness of the debate, even though the contribution itself is not a rational contribution, in the sense that its informational content contributes to the debate’s resolution. (Stanley 2015, 112)

This type of speech is, according to Stanley, structurally the opposite of demagoguery, that is, negative political propaganda.²³⁴ Both are nonrational contributions,²³⁵ and both have an effect on the subsequent reasonableness of a debate within a political culture. That is, both have an impact on the level of empathy within a community. But where demagoguery results in *less* empathy and *less* reasonableness, civic rhetoric *increases* empathy and therefore reasonableness.²³⁶ This *positive* propaganda is a type of “speech that uses ‘haranguing, cajoling,

²³³ I won’t say much more about Stanley’s analysis of negative propaganda here. Though, for interesting critiques that I am in broad agreement with, see McKinnon (2018), and Táíwò (2017).

²³⁴ Stanley presents positive propaganda as “structurally, precisely the opposite of demagoguery” (2015, 112) and makes this parallel explicit in later comments, where he says that “The distinction between civic rhetoric and demagoguery has solely to do with its ends, and not its form” (2018, 506).

²³⁵ Because of space limitations, I am unable to raise my concerns about this commitment of Stanley’s, which carves up speech acts in terms of them being ‘rational’ or ‘nonrational’ contributions.

²³⁶ Whether increasing reasonableness—especially via increasing empathy—is the proper goal is a question I put aside in this paper. Though, it should be noted that there may be significant drawbacks to dominant members attempting to themselves occupy the perspective of oppressed people, as the *aspiration* to empathy can itself backfire, as Iris Marion Young (1997) and others, have noted. One worry is the temptation to speak *for* another, and in doing so undermine their own (moral) authority. Thanks to Philip Yaure for pushing me on this.

threatening, or supplicating’ as a method to force a dominant majority to expand the domain of respect and empathy to include a persecuted and ignored minority (2015, 114).²³⁷

Beyond this functional similarity between demagoguery and civic rhetoric, Stanley also says there is a “structural problem in certain imperfectly realized liberal democracies that necessitates civic rhetoric” (ibid., 115). If a group lacks political power and has no say in a policy that affects them, their perspective is illegitimately left out, and—because of the prevalence of flawed ideologies—cannot easily be included through rational debate.²³⁸

There is a structural reason why this species of propaganda is a necessity in treating failures of liberal democracy. There are many times in which the perspectives of a group are invisible from the rest of the citizens. [...] In such a situation, there is no obvious deliberative way to make that group visible, no method that appeals to reason to bring members of that group into equal political standing. (2015, 114–115)

In short, inequality creates conditions where positive propaganda is necessary, in that it offers a path towards a more reasonable politics. It is in this way that Stanley describes the entire book as an argument for equality and says that much of Chapter Three “constitutes an argument that the aim of social movements should be civic rhetoric to expand empathy” (2018, 507). Drawing on an example from Du Bois, Stanley suggests civic rhetoric can play this role by employing liberal democratic ideals “against a certain understanding of their application”: freedom, solely for whites; or democracy, only for men.²³⁹

²³⁷ Stanley is here building off the ideas of Du Bois (1926) and Alain Locke (1928). “In this sense,” he says, “‘propaganda’ refers to a method of appealing to emotions to increase reasonableness” (2015, 111).

²³⁸ In a similar fashion, Elizabeth Anderson (2014b) notes that biases in moral thinking lead people holding powerful positions to “confuse their own power with moral authority,” and “to misread challenges to their orders from below as signs of vice.” In short, inequalities in power lead to arrogance and ignorance in the powerful.

²³⁹ For example, he suggests we understand Du Bois as “employing *the liberal democratic ideals themselves*, against a certain understanding of their application. His goal is to undermine a conception of liberal democracy that only extends freedom to whites. His method is to appeal to freedom itself, that liberal democratic ideal that is so cherished even among a nation in which it is restricted to whites. His rhetoric undermines an understanding of those ideals thus restricted, by calling attention to the fact that those ideals are deeply cherished among nonwhites as well. [...] “He is eliciting empathy by employing the problematically restricted ideals, and calling upon whites who cherish them to empathize with the plight of those who also cherish them, but to whom they have been consistently denied. The mechanism he is using is therefore a certain kind of undermining propaganda, one that targets the ideal

One example Stanley uses to explain this potential for positive propaganda comes from the Civil Rights Movement. The 1964 Selma to Montgomery March, Stanley says, “is a paradigm case of democratically acceptable propaganda: manipulation of the media to draw attention and empathy to the predicament of an otherwise invisible group.” What kind of manipulation is this? Stanley answers: King “manipulated white Southerners into revealing their hatred on national media, thereby turning the opinion of the country against them. (ibid., 113–114).

While this sounds to me like a roughly true description of historical events, this is where I will criticize Stanley’s application of positive propaganda to protest. First, it’s clear that Stanley’s main linguistic model of propaganda won’t apply, as it’s not obvious what utterance is expressed via the march, or how we could begin to demarcate the at-issue versus not-at-issue content.²⁴⁰ Second, this description centers the discussion on King’s manipulation of white Southerner’s cruelty—and the reactions of the white population more generally. While King and others certainly utilized this method to great effect, this ignores other core tenets of the movement that concerned the character of the resistor’s themselves, and how the anti-hierarchical and democratic ethos of nonviolent direct action are virtues in and of themselves.²⁴¹ Stanley’s model focuses mainly on the *audience* of the protest rather than the protesters themselves, and I believe there is a significant cost to this.

freedom just for whites. [...] In this way, the liberal concepts can be used against restricted understandings of their proper application” (2015, 116–177)

²⁴⁰ The content of utterances at protests are, of course, relevant, in part because they can help up distinguish egalitarian protest from bigoted protest, and so could perhaps play this role. Rather than take this route, I’m interested here in how content contributes to the pragmatic structure—who it’s calling on and how. That is, how the speech acts of protest function as a second-personal transactions. For more on this, see (Kukla and Lance, 2009; Lance and Kukla, 2013; Herbert and Kukla, 2016).

²⁴¹ See King (1963, especially Chapter Two) for this argument.

As I will suggest, then, there is more that distinguishes positive and negative propaganda at the structural level than their effects—at least when discussing *protest* as a form of positive propaganda.²⁴² To make this more explicit, and to demonstrate what is left out, I now turn to a contemporary example: BLM.

2.1.1. *Black Lives Matter as positive propaganda*

A lot has been written about the Movement for Black Lives in general, and Black Lives Matter protests in particular,²⁴³ and to be clear my aim here isn't to in any way give a definitive analysis of an ongoing phenomena. Rather, what I aim to show is that analyses of protest that focus at the level of expressed content—like Stanley's mode of positive propaganda—encourage obscure important aspects of the pragmatic of protests. Namely, it obscures protest's function to foreground the moral authority of the protester in a way that challenges the unjust authority of powerful. Discussions of Black Lives Matter protest are an illuminating example of this in part because so much attention has been given to the phrase that serves as both the group's name, and their main slogan. To see why, consider two fairly popular—and in my eyes, representative and revealing—discussions of BLM protest.

The first comes from comedian and activist Franchesca Ramsey's video titled "4 Black Lives Matter Myths Debunked" (2016). There, Ramsey makes the point that the phrase "Black Lives Matter" is not a racist statement directed against non-Black people and should be interpreted as saying "Black Lives *Should* Matter." She says:

²⁴² I'm not sure we *should* think of protest as a form of propaganda. But I'm interested in following Stanley down this path, if only to see what we might learn.

²⁴³ See Taylor (2016), and Lebron (2017) for examples of some recent academic analysis.

This movement isn't saying black lives matter more than anyone else's. It's saying that Black lives should matter, but the way that our justice system, our media, and our police have been operating suggests that they do not. (ibid.)

This tries to clarify a surprisingly common (among whites) misconception about the origins and meaning of these protests. And it does so by uncovering the meaning—that is, semantic content—of the slogan.

Law School professor Patricia Leary takes this approach a bit further in her viral letter written in response to anonymous student complaints about her wearing a BLM pin to class. Leary carefully analyzes the claims of her student detractors and points out that their argument rests on a false premise, namely, that “there is an invisible ‘only’ in front of the words “Black Lives Matter.””²⁴⁴ As Leary goes on to say, while this assumption is false, she suggests that:

there *are* some implicit words that precede ‘black lives matter,’ and they go something like this: ‘Because of the brutalizing and killing of black people at the hands of the police and the indifference of society in general and the criminal justice system in particular, it is important that we say...’

This, as she points out however, doesn't nicely fit on a shirt.

These analyses suggest that the phrase “Black Lives Matter,” and the protests of which it symbolizes, fit the model of positive propaganda that Stanley articulates and may be productively analyzed as such. Through their provocative slogan, BLM deploys nonrational means of persuasion that aim to extend empathy to an oppressed group. We can see this because:

- (1) the message “Black Lives Matter” embodies a cherished moral ideal—moral value, civic equality;
- (2) by extending this ideal to group that has been unjustly excluded from the dominant interpretation of that ideal this reveals how it is, in practice, restricted to whites;
- (3) in doing so, it undermines the existing, restricted ideal of moral value; and
- (4) this has the effect—or aims—of increasing empathy for the excluded group.

²⁴⁴ See Jaschik (2016) for the letter and context.

We see this, moreover, when we engage in the kind of excavation of hidden meaning that Stanley’s account encourages, searching for the unarticulated constituents that reveal the true power of the slogan.

A similar analysis could be provided for the counter-slogan “All Lives Matter” that explains how, given the context in which this phrase emerged—that is, in direct response to Black Lives Matter protests—the expressed content of “All Lives Matter” contains implicit associations that function to silence black protest, and ultimately reduce empathy. A worthy political ideal—moral and legal equality—is appealed to in a fashion that in fact serves to undermine that very ideal, by presenting Black Lives Matter protestors as racially partisan, and therefore ignorable. And, without that context, it’s so abstract to be meaningless. Something Paul C. Taylor (quoted in Talisse 2017) recently called “the race to the trivial.”

2.1.2. *What’s missing*

While I don’t necessarily disagree with the analyses discussed above,²⁴⁵ I do worry that this type of analysis, with its focus on semantic content—including implicit or presupposed content—comes at a cost. When considering slogans as they’re used in political protest, I agree with Elliott Colla (2013, 45), who argues that “the context of performance demands that we consider slogans not just in terms of semantic meaning or as discursive genre, but also as embodied actions taking place in particular situations.”²⁴⁶ He elaborates:

²⁴⁵ And in the case of ‘All Lives Matter’ I think this analysis is illuminating.

²⁴⁶ Colla, focuses on the further non-semantic features of the protest slogans of 2011 Egyptian revolution. While I don’t want to suggest all protest are the same, I think this lesson is generalizable. Furthermore, in considering the minor literature on slogans, Colla notes that “the tendency is to treat the slogan primarily as a semantic text” (2013, 38). And it is exactly this that I want to move away from.

slogans and chants may be composed of words, but they are not merely linguistic texts. Rather, they are part of something that can only be called public performance. Thus the meaning of revolutionary slogans cannot be reduced to any semantic content, nor can it be translated as a set of linguistic statements. (Colla 2013, 38)

This is good reminder about the context and performative aspect of protest—most notably, how protest slogans are (usually) spoken by a group of people. When we lose sight of this and ignore these elements, and instead stick to the semantics, we tend to push the analysis and subsequent discussion in a specific direction. Namely, this presents the speech acts of protests as though they are *moves in a debate*. That they, and the counter-speech they generate, are competing claims deserving of equal attention. This is, I will argue, an inadequate approach to take for protest. As Bernard Boxill (1976) states:

Typically, people protest when the time for argument and persuasion is past. They insist, as Du Bois put it, that the claim they protest is “an outrageous falsehood,” and that it would be demeaning to argue and cajole for what is so plain. Responding to a newspaper article that claimed “The Negro” was “Not a Man,” Frederick Douglass disdainfully declared, “I cannot, however, argue, I must assert.”

As Boxill might put it, treating “Black Lives Matter” and “All Lives Matter” as claims competing in a debate obscures some of the central features that are inherent in protest, namely, that moral protest involves a *demand*. A type of demand, moreover, that asserts the protester’s moral entitlement to make such a demand—that is, their authority *to demand*.

While there are claims being put forward by BLM that ought to be evaluated and debated—e.g., their policy proposals—the act of protest itself is not something that deserves to be argued over.²⁴⁷ The call of “Black Lives Matter” is not put forth as a claim to be contested, as if it were a premise in an argument in a seminar room. Recognizing it *as protest* means

²⁴⁷ And compare this focus on unarticulated constituents to the comments made by Patrisse Khan-Cullors, quoted in the epigraph to this chapter. This, of course, is not to say that the founders of Black Lives Matter did not think long and hard about their choice of phrasing; they did. For more, see Khan-Cullors, and bandele (2018).

recognizing it as a speech act *to be heard*. As I see it, this is a difference in the *uptake* the speech act aims at, and this is obscured when we focus on the proposition being expressed. Speech acts like assertions seek uptake “in the form of agreement or rational challenge from others” (Kukla 2014, 452). But *protest* calls for a different type of uptake, and this is closer to the *imperatives* or *orders* than assertions, because they call “for is obligations as their upshot” (Lance and Kukla 2013, 459). Recognizing as much is the first step towards seeing the speech act of protest for what it is.

Beyond this difference in uptake, the entitlement conditions that the speech act of protest presumes are another area worthy of attention. That is, one contribution that protest makes to a political culture concerns not simply what’s being said, but *who gets to say it*. From what *perspective*, or social-location is the protest being asserted?²⁴⁸ In their work, Herbert and Kukla (2015) point out the existence of “community-specific speech,” which has both community-specific input and uptake. On the input side, they note how “speech acts that have community-specific inputs are of a sort that are felicitous only when performed by insiders” (580). I want to show how this is an important feature of the input of protest,²⁴⁹ one that I explain in terms of collective authority. Recall my description of the authority at issue in collective speech acts from Chapter Three:

Collective authority: Through repetition and endorsement, individually non-authoritative speech is made authoritative by signaling collective support and solidarity. A group of speakers have their authority secured through a repetitive process where each new utterance adds to the strength of the overall practice.

²⁴⁸ As Khan-Cullors and bandele (2018) note about BLM: “We have centered and amplified the voices of those not only made most vulnerable but most unheard, even as they are on the front lines at every hour and in every space: Black women—all Black women.”

²⁴⁹ Though, because it is (most often) directed at outsiders, it is not itself community-specific speech in Herbert and Kukla’s (2015) sense.

To build towards a conception of protest that employs a similar type of collective authority, I first turn to Elizabeth Anderson for a helpful analysis of the function of social movements.

3. THE MORAL-EPISTEMOLOGICAL FUNCTION OF SOCIAL MOVEMENTS

In her analysis of the British Abolition movement, Elizabeth Anderson notes that social movements tend to do three things to correct for the biases of the powerful:

- (1) They *inform* the powerful of the needs and interests of the less powerful.
- (2) They *express* what is required to respect these needs and interests *as claims or demands* on the powerful.
- (3) they *enable* the less powerful to *display their worthiness*, so that they can assume some *moral authority* to contest the counterclaims of the powerful, and put authority behind their own claims. (2014b, 8, emphasis in original)

It is features like these, which are informed by elements such as speaker, audience, context, etc., and cannot be captured in impersonal propositional terms that I believe an analysis of protests must address.

The protest of social movements, according to Anderson, functions to reject the authority of the unjust norms of the dominant. In their place, they foreground the collective authority of the protesters themselves. They position their own voices as voices in need of being heard. And this act, when performed by oppressed persons, directly threatens an unjust hierarchy. Note, moreover, how this is in some ways aside from the specific content they express. As Judith Butler remarks:

To speak truth to power is not fundamentally an individual act. Before we ask what it means to speak truth to power, we have to ask who can speak. Sometimes the very presence of those who are supposed to remain mute in public discourse breaks through that structure. [...] So although we can think about parliamentary assemblies as part of democracy, so too can we understand the extraparliamentary power of assemblies to alter

the public understanding of who the people are. Especially when those appear who are not supposed to appear. (2017)²⁵⁰

Failing to recognize this aspect of protest and focusing too closely on content, then, tempts us to mischaracterize protest as argument. This threatens to make us lose sight of the concrete social context that produced the protests, along with its embodied performance. We should, in contrast, see protest as type of speech that foregrounds the collective authority of its speakers. And so, as instances of positive propaganda—where they aim to or manage to increase the empathy of their audience—it is through a distinct and more direct means than those highlighted by Stanley.

To offer a clearer picture of what I mean by the ‘moral authority of the protester,’ I first turn to how a similar concept is discussed in a different context: the (meta)ethics of moral demands. Then I return to the main topic of this dissertation: the pragmatics of hate speech. By looking to these two areas, we will find a clearer and richer picture of the authority—the entitlements—protesters are calling upon in their (speech) acts.

3.1. Egalitarianism, Protest, and Second-Personal Calls

To better approach the role of authority in protest, it’s useful to turn to a branch of egalitarianism known as “relational egalitarianism.”²⁵¹ What makes relational egalitarianism

²⁵⁰ Butler (2017) continues: “Although demonstrations and assemblies are often not enough to produce radical change, they do alter our perceptions about who the people are, and they assert fundamental freedoms that belong to bodies in their plurality. There can be no democracy without freedom of assembly, and there can be no assembly without the freedom to move and gather. When the undocumented assemble, or when those who have suffered eviction assemble, or those who suffer unemployment or drastic cuts in their retirement, they assert themselves into the imagery and the discourse that gives us a sense of who the people are or should be. Of course, they make specific demands, but assembly is also a way of making a demand with the body, a corporeal claim to public space and a public demand to political powers.”

²⁵¹ For more on this form of egalitarianism, see Anderson (1999), Wolff (1998; 2019), Scheffler (2010), and Fourie, Schuppert, and Wallimann-Helmer (2015).

distinctive is its central concern with social hierarchies and interpersonal power.²⁵² The central questions for many relational egalitarians then become: what is it that we *owe* one another, and what can we *demand* from each other?²⁵³ In being so focused, relational egalitarians highlight the “second-personal” dimensions of many core ethical concepts, like rights, duties, and justice.²⁵⁴ On this view, complaints of injustice are best seen as demands, which casts them as second-personal utterances rather than impersonal expressions of propositions. A second-personal utterance is a speech act that is directed at and calls upon a second person, a “you,” to give it specific uptake.²⁵⁵ And so, a moral claim is a performative utterance, where one makes a claim—a demand—on another, and at the same time asserts their entitlement to do so. “A claim of justice,” Anderson (2012, 3) says, “is essentially expressible as a *demand* that a person makes *on* an agent whom the speaker holds accountable.” Therefore, on this reading, it matters to whom one makes a claim on, and whether the speaker is properly entitled with the proper moral authority to do so. Different kinds of calls create different kinds of outputs. Second-personal calls create distinct, relational statuses between the speaker and her target audience.

It’s worthwhile to note that, for some philosophers in this camp, the inspiration for the theory comes directly from egalitarian social movements. For instance, both Anderson and Iris Marion Young take such social movements as the Civil Rights movement, LGBTQ+ rights advocates, and more, to be crucial for both our theorizing about concepts like justice, and to be

²⁵² This is opposed to the focus on material inequality that resource-egalitarians about distributive justice adopt. Though, it should be noted that both forms of egalitarianism are concerned with various inequalities of different sorts—social, material, welfare, etc.—the difference lies mainly in what is given greater explanatory power, along with distinct metaethical commitments. See Anderson (2010; 2012).

²⁵³ The notions of second-personality found in this literature, including the centrality of demands, is mainly inspired by Darwall (2006).

²⁵⁴ Anderson (2012) makes this most explicit.

²⁵⁵ For more on the second-personal aspect of calls, see Kukla and Lance (2009); and Lance and Kukla (2013).

themselves a core engine of moral progress.²⁵⁶ As Anderson (2014a, 260) says, social movements are “the source of egalitarian ideas.”

This appreciation of the second-personal nature of moral claims recognizes that protest and demands are distinct from mere moral arguments, which are indifferent to elements like speaker, audience, context, etc. Pure moral arguments—the type we might read in a philosophy paper or discuss in an ethics classroom—are often expressed in third-personal language to emphasize their presumed universality. There is a difference, therefore, between the *argument* that all moral agents should refrain from hurting other sentient beings because of the badness of pain, and the second-personal *demand* that you stop stepping on my toe.²⁵⁷ The latter, but not the former, highlights and grounds out in my authority to make a claim upon you and hold you accountable. It is this (meta-ethical) difference in address, that places protests in between the poles of “pure moral argument” on the one hand, and “riots, war, and other violent acts” on the other (Anderson, 2014b, 9). The broader point is that understanding moral claim-making in the real world requires taking stock of these broader contextual features that give moral life its richness and specificity.

One element this attention to context reveals, then, is the invocation of the moral status of the person making a moral demand. Therefore, in the case of protest, the status and position of the protester is essential to fully understanding the act being performed. In protesting, one does more than express dissatisfaction with the status quo. They position this complaint as originating from a specific social location. Moreover, as we’ve seen, protest is fundamentally *not* an

²⁵⁶ See Anderson (2014a), and (2014b), and Young (2011).

²⁵⁷ To borrow one of Darwall’s (2006) favorite example.

individual act, but is rather performed as part of a larger group-action. This is evident when we look at the embodied practice of the speech acts of chanting slogans, As Colla notes:

Slogans are chanted, shouted and sung by embodied people moving, often in coordinated ways, in and through public spaces; and 2) these movements and actions are not mere context for the production of slogan meaning, but are part of the text itself. (2013, 38)

It is in this ways that the coordinated, collective speech acts of protesters call upon *their* moral standing and situated knowledge, issuing second-personal calls with distinct entitlements.²⁵⁸ That is, they put some distinctly *collective* authority behind their claims, and we must keep this in view.

To further develop this notion of authority, I'll now return to the topic of subordinating speech. This will help to fill in additional details concerning the role of context and social position, specifically with regards to how these inform speaker authority and shape the pragmatic force of speech acts. This helpful because, moral demands like “get off my toe” require only a general entitlement which most of us share, but, as we've seen many speech acts require a *particular* type of authority to be successfully carried out.

3.2. *Subordinating Speech and Collective Authority*

Recall why authority seems necessary to account for the force of hate speech. Once we consider what *act* is being done when one person hurls hate speech at another, we see these as *degrading* and *subordinating* acts. And so, to account for the pragmatic force of such speech acts, we're drawn to the idea that hate speakers draw on some form of authority to perform these

²⁵⁸ I believe reclamation projects are a helpful example of this, in part because reclamation depends on centering the perspective (and authority) of members of the group targeted by the contested term. That is, we cannot lose sight of the 'we' in “we're here, we're queer, get used to it.” Note, though, that reclamation is far from straightforward. For discussion of these issues, see Herbert (2015); and Tirrell (1999).

acts with their words. And as we've seen, subordinating speech generally relies on a *dominating relation* to perform its characteristic function of subordinating its targets.

Subordinating speech is therefore *asymmetrical speech*, and speaker and target are not equally situated in regards to their speech capacity—it's just not the case that the target of hate speech can just turn on their assailants and return fire with fire.²⁵⁹ But, as I've argued throughout, this authority is not essentially tied to *formal positions* of authority, and is achievable in a number of distinct ways. The authority at issue here relies on a richly contextual network of features that grant some, but not all, distinct normative powers in particular situations. Informal distinctions of power and privilege along lines of race, gender, ability, etc., play a large role in distributing this authority, significantly affecting the type of speech acts available to different speakers.²⁶⁰ This reveals how authority can be thoroughly contextual and interpersonal.

As I argued over the last two chapters, an important type of subordinating authority is *collective authority*. This is generated through repeated use of the same or similar utterances by a group of speakers. Where each individual utterance would seem to lack authority when considered in isolation, it is by noting the pattern of repetition and amplification that we can understand these as authoritative—and subordinating—speech acts. Through this process, individual speakers fall from view, and instead, a somewhat faceless mass of speakers join to produce speech with a distinct, and stronger, pragmatic force.²⁶¹

Protest, I believe, are another interesting instance of this. Like online abuse, the speech acts of protest gain a significance and strength that is incomprehensible when considered as simply originating from individual speakers. And to fully understand the speech acts of protest,

²⁵⁹ And this shows the limits of the 'more speech' response to hate speech.

²⁶⁰ For a structural account of the harms of injustice in speech in terms of speech capacity, see Ayala (2016).

²⁶¹ See Chapters Three and Four.

we must attend to this aspect of its pragmatic output. Colla (2013) describes this feature of a protest slogan in terms of how it “is intended to circulate as an authorless text.” He goes on to say that:

one index of a slogan’s power is the degree to which it can detach itself from the specific conditions of its initial composition, and the degree to which it circulates as if it were the anonymous expression of a collective will. (Colla 2013, 38)

Of course, by anonymous, this does not mean that the speech acts of protest are identity-less. The social location of those protesting is obviously significant. But, as this shows, it’s the identity of the *group* as a whole that’s relevant. This is, I believe, similar to how the ‘real’ identity of online trolls falls from view when their speech acts are understood as part of a larger, coordinated group activity. And, at least in one sense, it is *through speech* that this group comes into being, and acts as a collective. This is, in part, what these collective speech acts do, as their pragmatic function. They create or solidify a group identity. “Slogans are performatives,” Colla says, “in the sense they are deliberate compositions intended not so much to reflect collective will *but to create it*” (2013, 38, emphasis added).

In the final section of this chapter, I return to the topic of how differences at the pragmatic level of speaker authority and uptake distinguish protest from subordinating speech, like propaganda. To help illustrate my claims, I first go over the idea of protest as *provocation* in order to show how putting the speech acts protest in their proper context, which often means receive some counter-protest speech in response, clarifies both speech acts.

4. PROTEST AS PROVOCATION

Thinking of protest as *provocation* helpfully demonstrates an important interplay between protest and (a subset of) its audience. Booker T. Washington's dismissal of protest out of prudence was partly based on this. Washington claimed that one danger of protest lies in its potential for provocation. "A provocation arouses an individual's resentment," Boxill (1976) notes, "because it challenges his moral claim to a status he enjoys and wants to preserve, thus black protest would have challenged the white South's justification of the superior status it claimed." The concern here is that protest arouses resentment because it constitutes a challenge to its targets, in part because it questioned their status. White Southerners would not accept the (equal) entitlement Black protesters claimed *in protesting*, and thus would react with hostility.

Kate Manne raises a similar worry in her analysis of the violence and scorn directed at BLM protesters in Ferguson. She writes:

The humanist line on Ferguson [i.e., that the violence and hatred is sourced in white American's inability to see Black Americans as fully human] fails to explain what seems to provoke the aggression—namely, acts of political and personal defiance, which only people can demonstrate. Moreover, it is hardly surprising historically subordinated people should be perceived in this way when they try to assert themselves around, or over, dominant group members. They are liable to be perceived as belligerent, 'uppity,' insubordinate or out of order. (2014)

As this shows, protest, because it both presumes a certain entitlement on the part of the speaker, and makes a specific claim on its target, can serve to trigger hostile and resentful reactions from those on whom the claim is being made. After all, as we saw, protest functions as a demand. But in a hierarchy, not everyone can (successfully) make a demand upon another, as demands are also asymmetrical speech acts in this context.

So, in addition to the presumption of authority—the moral entitlement *of the protester*—protest also implicates the *status of its target*—the group to whom the protest is directed. In this

way, it is like subordinating speech, in that it depends on the dynamics of power in the entire relation between speaker and target. Protest makes vivid not only the social position of the protesters themselves—their collective authority—but also the social position of their audience, that is, the more powerful. It challenges the justification for their superior status, which means it is often interpreted as threatening—because, in important ways, it is. And yet, of course, it is the denial of equality and justice that make protest necessary.

So, under conditions of oppression, “acts of personal and political defiance” on the part of the oppressed, like protest, are in and of themselves a challenge to the status quo, and it is the broader political and contextual features that make this vivid, rather than any particular aspect of the protest’s content. Putting protest in context, seeing it as provocation that can lead to a range of reactions reveals this aspect more fully.

4.1. A Strategic Objection

This analysis, however, leads us to an important complicating factor, namely, that conditions of oppression lead many members of dominant social groups to dismiss oppressed people as sites of moral and epistemic authority. And hostile threats to the status quo are interpreted as proof of their unequal status. That is, anti-Black racism makes it difficult for people in a white-supremacist society to accept Black protesters as moral authorities in need of being listened to. To take one example, Shree Paradkar (2017) notes this in writing about the 1992 Yonge St riots in Toronto. “Blacks who protest violently are thugs,” she writes, while “whites who do so have a righteous anger.”²⁶² We see this time and again.

²⁶² She goes on to add: “The police continue to allege criminality by the victim and a society that prides itself on its decency silently shakes its head—not at the injustice but at the tactics of those who protest that injustice, the

One possible implication might be then, where there exists entrenched racisms and other forms of marginalization, talk of the moral authority of the protester is a luxury reserved for the more privileged. It's no accident that Anderson's analysis of British abolitionism focuses mainly on the white Britons who opposed the slave trade but weren't themselves at risk of enslavement. This may lead one to suggest that, given this racist resistance and the tendency to see protest as provocation, it might be a strategic move to focus on the *content* of the claims being made rather than the people. If Black protesters are more likely to be perceived as violent criminals rather than authoritative voices on injustice—at least for white audiences—then maybe impersonal arguments are more effective.

This is a serious concern, for both activists and for how protest ought to be analyzed and discussed. And I can offer only a partial response. One thing to note, though, is that a similar danger lurks when looking at protest in terms of presenting an argument, so perhaps this is no refuge. We see this when the particular claims being made by *certain* speakers—e.g., people of color—are subject to heightened scrutiny. Melissa Harris-Perry (2011) notes that when she and others write on racism in the United States, they tend to “encounter a few common discursive strategies that are meant to discredit our perspectives.” She writes that some common tropes emerge in response to an analysis that suggests racism might be the cause of someone's actions, for example, the demand to “prove it,” and the related skepticism implicit in questions of “who made you an expert?”²⁶³ Together these tropes demonstrate the raised standards many claims of racial bias face, especially as they come from people of color.²⁶⁴

injustice of their deaths at the hands of their defenders.” This is, of course, a familiar phenomenon, and I only use this example because it concerns Toronto, where I am from.

²⁶³ She also notes the “I have black friends” deflection, though this is less relevant to my case here.

²⁶⁴ See Fricker (2007) for a discussion of “epistemic injustice” that aims to make sense of these dynamics.

And while there is no doubt people of color encounter these ‘strategies’ more severely than whites, part of Harris-Perry’s point seems to be that in a racist society, the topic of racism itself invites frustrating skepticisms. As she notes, one strategy is simply to “scorn the study of race as an illegitimate intellectual pursuit.” As I see it, this heightened scrutiny relates in some ways to Kristie Dotson’s discussion of “risky” and “unsafe content” that can lead a speaker to silence herself, to pre-emptively avoid such scrutiny (Dotson, 2011). So, where racial injustice is present, even when the focus becomes content, it is similarly likely for some—typically white—audiences to dismiss the concern being raised. And so, a retreat to the (disembodied) content being expressed offers no safe refuge from anti-Black racism.

Provocation, in other words, is inevitable when one calls out injustice. And despite the criticisms I have raised above, Stanley has a clear grasp of this sad feature of political speech. “A salient feature of many paradigm cases of propaganda,” he says, “is that it is speech that owes its efficacy in ending rational debate not to its settling of the question, but rather to its erosion of second-personal ideals like reasonableness” (2015, 121). And so, we can see how many paradigm cases of propaganda would emerge as attempts to renounce the efforts of social movements. The ‘all lives matter’ response to the call of “Black lives matter” demonstrates this very efficiently. Protest invokes a demand—which presupposes the moral authority to do so—and counter-protest propaganda looks to deny this entitlement.²⁶⁵ But as I have argued, an analysis that focuses at the level of content misconstrues the way these speech acts differ. Instead, it helps to see these acts in conversation with each other to see this crucial difference in uptake each aims at.

²⁶⁵ And so, both may be examples of what Lance and Kukla dub an “activist speech act,” that is, “one that does not merely draw upon but in fact functions to subvert or reconfigure the normative context in which it operates” (2013, 473).

Indeed, this interplay works to reveal the element of protest that I've argued Stanley's account ignores. That is, while Stanley sees civic rhetoric and demagoguery—positive and negative propaganda—as structurally parallel, his analysis of both focuses most extensively on the effect each has on its *audience*. Demagoguery reduces empathy; civic rhetoric increases it. But the role of the speaker is obscured, as is the relationship between speaker(s) and audience.

This is inappropriate, I argue, in the case of protest, which is a speech act that fundamentally aims to reveal the moral authority of the protesters, and in doing so relies on and makes us of a distinct form of collective authority. And it is in this way that protest may be an instance of positive propaganda. Not because it engages in clever (linguistic) manipulation, but because, through the act of protest oppressed people claim what they are properly entitled to, and it is this assertion of authority in the face of injustice that effectively serves undermines unjust domination.

5. CONCLUSION

In this chapter I examined what kind of speech act protest is from a pragmatic point of view. To do so, I questioned Jason Stanley's claim that protest may be a form of *positive propaganda*. However, as I argued, his model of propaganda is an ill fit for paradigmatic types of egalitarian protest, like BLM protests, because of its single-minded focus of the effects of protest on the audience. Instead, I called attention to the speakers' actions—the imposition of a *demand* on the unjustly dominant—and argued that understanding what makes this act possible requires different tools. By considering the distinct uptake protest aims at, along with the specific type of *collective* authority this act exerts, I suggested that the emancipatory potential of protest

consists fundamentally in its capacity to wield this authority. We see this, I argued, not by questioning what protest slogans 'mean' outside of their context, but by seeing the people engaging in these acts as embodied people, joining together to perform acts that would not be possible separately.

It is in this way that the speech acts of protests are relevantly similar to some subordinating speech acts. Both are only properly understood by attending to the specific type of authority being invoked, and both rely on acts of solidarity to generate this specific type of collective authority. Looking at either isolated from its context of use offers only a poor understanding of the speech act being done. But by attending to the dynamic context of use, our understanding of each is enlightened.

REFERENCES

- Adler, Simon. 2018. "Post No Evil." *Radiolab*. WNYC Studios. August 17.
<https://www.wnycstudios.org/story/post-no-evil>
- Althusser, Louis. 1971. "Ideology and Ideological State Apparatuses." *Lenin and Philosophy and Other Essays*. Translated by Ben Brewster. Monthly Review Press.
- Anderson, Elizabeth. 1999. "What is the Point of Equality?" *Ethics* 109, 287–337.
- . 2010a. *The Imperative of Integration*. Princeton: Princeton University Press.
- . 2010b. "Justifying the Capabilities Approach to Justice." In *Measuring Justice: Primary Goods and Capabilities*, edited by Harry Brighouse and Ingrid Robeyns, 81–100. Cambridge: Cambridge University Press.
- . 2012a. "Equality." In *Oxford Handbook in Political Philosophy*, edited by David Estlund, 40–57. Oxford: Oxford University Press.
- . 2012b. "The Fundamental Disagreement between Luck Egalitarians and Relational Egalitarians." *Canadian Journal of Philosophy*, Supplementary Volume 36: 1–23.
- . 2014a. "A World Turned Upside Down: Social Hierarchies and a New History of Egalitarianism." *Juncture* 20 (4): 258–261.
- . 2014b. "Social Movements, Experiments in Living, and Moral Progress: Case Studies from Britain's Abolition of Slavery." Paper presented at the University of Kansas as part of the Lindley Lecture Series, Lawrence, Kansas. February 11.
- Anderson, Luvell. [Forthcoming]. "Notorious Thugs." Unpublished manuscript.
- Anderson, Luvell and Ernie Lepore. 2013. "Slurring Words." *Noûs* 47: 25–48.
- Ashwell, Lauren. 2016. "Gendered Slurs." *Social Theory and Practice* 42 (2): 228–39.
- Austin, J. L. 1962. *How to do Things with Words*. Cambridge: Harvard University Press.
- . 1979. "Performative Utterances." In *Philosophical Papers*, 3rd Edition, edited by J. O. Urmson and G. J. Warnock, 233–252. Oxford: Clarendon Press.
- Ayala, Saray. 2014. "Explaining Injustice in Speech: Individualistic vs. Structural Explanation." Unpublished Manuscript.
- . 2016. "Speech Affordances: A Structural Take on How Much we can do with our Words." *European Journal of Philosophy* 24 (4): 879–891.
- Ayala, Saray, and Nadya Vasilyeva. 2014. "Allowing Harmful Speech: Responsibility for Silence." Unpublished Manuscript.
- Bakalar, Chloé. 2018. "Programming Deliberation: The Ethical and Political Implications of Automated Content Moderation Algorithms." Paper presented at the Georgetown Political Theory Workshop. February 22.
- Barbaro, Michael. 2017. "'The Daily': Alt-Right and the Internet." *New York Times*. August 18.
<https://www.nytimes.com/2017/08/18/podcasts/the-daily/alt-right-social-media.html>
- Barney, Rachel. 2016. "[Aristotle], On Trolling." *Journal of the American Philosophical Association* 2 (2): 193–195.
- Bauer, Nancy. 2015. *How to Do Things with Pornography*. Cambridge: Harvard University Press.
- Bartlett, Jamie. 2014. *The Dark Web: Inside the Digital Underworld*. New York: Melville House.
- BBC News. 2017. "Fake Obama Created Using AI Video Tool." *YouTube*. July 19.
<https://www.youtube.com/watch?v=AmUC4m6w1wo>
- Beauhamp, Zack. 2019. "Our Incel Problem." *Vox*. April 23. <https://www.vox.com/the-highlight/2019/4/16/18287446/incel-definition-reddit>

- Bell, W. Kumau, and Hari Kondabolu. 2017. "Patrisse Khan-Collurs on Black Lives Matter and Resistance under 45." *Politically Reactive*, season 2, episode 1. <https://www.politicallyreactive.com>.
- Belmont, Veronica. 2018. "Cloak of Invisibility." IRL: Online Life is Real Life, a podcast by Mozilla. February 18. <https://irlpodcast.org/episode11/>
- Berdini, Federica. 2013. "Speech Acts and Normativity: A Plea for Inferentialism." *Esercizi Filosofici* 8 (2): 71–88.
- Bernstein, Joseph. 2015. "In 2015, The Dark Forces of the Internet Became a Counterculture." BuzzFeed. December 23. https://www.buzzfeed.com/josephbernstein/in-2015-the-dark-forces-of-the-internet-became-a-countercult?utm_term=jcmLwKXIW#.hbPV3edYy
- . 2018. "The Unsatisfying Truth About Hateful Online Rhetoric and Violence." BuzzFeed News. November 23. <https://www.buzzfeednews.com/article/josephbernstein/does-hateful-speech-lead-to-violence>
- Bolinger, Renée Jorgensen. 2017. "The Pragmatics of Slurs." *Noûs* 51 (3): 439–462.
- Boxill, Bernard. 1976. "Self-Respect and Protest." *Philosophy and Public Affairs* 6: 58–69.
- Brandom, Robert. 1998. *Making it Explicit*. Cambridge: Harvard University Press.
- Brennan, Jason. 2017. "Propaganda About Propaganda." *Critical Review* 29 (1): 34–48.
- Brown, Alex. 2017. *Hate Speech Law: A Philosophical Examination*. London: Routledge.
- Butler, Judith. 1997. *Excitable Speech: A Politics of the Performative*. New York: Routledge.
- . 2017 "Reflections on Trump." *Cultural Anthropology* January 18, <https://culanth.org/fieldsights/1032-reflections-on-trump>.
- Camp, Elisabeth. 2013. "Slurring Perspectives." *Analytic Philosophy* 54 (3): 330–349.
- Cassidy, John. 2018. "A Racist in the Oval Office." *The New Yorker*. January 12. <https://www.newyorker.com/news/our-columnists/trump-shithole-comment-racist-in-the-oval-office>
- Catterall, Helen T., ed. 1926. *Judicial Cases Concerning Slavery and the Negro*. Washington, D.C.: Carnegie Institute.
- Cavell, Stanley. 2005. "Performative and Passionate Utterance." In *Philosophy the Day After Tomorrow*, 155–191. Cambridge: Harvard University Press.
- Cecco, Leyland. 2018. "Canada Mosque Shooter Says He was Motivated by Trudeau Welcoming Refugees." *The Guardian*. April 13. <https://www.theguardian.com/world/2018/apr/13/canada-mosque-shooter-alexandre-bissonnette-trudeau-trump-refugees-travel-ban>.
- Chait, Jonathan. 2017. "How 'Neoliberalism' Became the Left's Favorite Insult of Liberals." *New York Magazine*. July 16. <http://nymag.com/daily/intelligencer/2017/07/how-neoliberalism-became-the-lefts-favorite-insult.html>
- Cherry, Myisha. 2015a. "Twitter Trolls and the Refusal to be Silenced." In *The Real World Reader: A Rhetorical Reader for Writers*, edited by James S. Miller, 221–225. [New York: Oxford University Press].
- Crichton, Michael. 2003. "Aliens cause Global Warming." Speech delivered at the California Institute of Technology. January 17.
- Colla, Elliott. 2013. "In Praise of Insult: Slogan Genres, Slogan Repertoires and Innovation." *Review of Middle East Studies* 47 (1): 37–48.
- Collins, Patricia Hill. 1998. "It's All In the Family: Intersections of Gender, Race and Nation." *Hypatia* 13 (3): 62–82.
- Collins, Patricia Hill, et al., 2002. "Symposium on West and Fenstermaker's 'Doing Difference,'" in *Doing Gender, Doing Difference: Inequality, Power, and Institutional Change*, edited by Sarah Fenstermaker and Candace West, 81–94. New York: Routledge.
- ContraPoints. 2018. "Incels." *YouTube*. August 17. <https://www.youtube.com/watch?v=fD2briZ6fB0>
- Cudd, Ann E. 2006. *Analyzing Oppression*. New York: Oxford University Press.

- Darwall, Stephen. 2006. *The Second-Person Standpoint: Morality, Respect, and Accountability*. Cambridge: Harvard University Press.
- Delgado, Richard. 1993. "Words that Wound: A Tort Action for Racial Insults, Epithets, and Name Calling." In *Words that Wound: Critical Race Theory, Assaultive Speech and the First Amendment*, edited by Mari Matsuda, Charles Lawrence III, Richard Delgado, and Kimberlé Crenshaw, 89–110. Boulder: Westview Press.
- Delgado, Richard, and Jean Stefancic. 2004. *Understanding Words that Wound*. Boulder: Westview Press.
- Du Bois. 1926. "Criteria of Negro Art." *The Crisis*, Volume 32, October, 290–297.
- Dotson, Kristie. 2011. "Tracking Epistemic Violence, Tracking Practices of Silencing." *Hypatia* 26 (2): 236–257.
- Ellul, Jacques. 1973. *Propaganda: The Formation of Men's Attitudes*, vintage books edition. New York: Vintage Books.
- Fanon, Frantz. 2008. *Black Skin, White Masks*. Translated by Charles Lam Markmann. London: Pluto Press.
- Fine, Cordelia. 2010. *Delusions of Gender: How Our Minds, Society, and Neurosexism Create Difference*. New York: W. W. Norton & Company.
- Fisher, Max, and Amanda Taub. 2019. "Social Media has a Mob Violence Problem. Could Soccer Hooliganism Prevention Offer a Model for Solving it?" *New York Times, The Interpreter*, June 6. https://static.nytimes.com/email-content/INT_14028.html
- Fourie, Carina, Fabian Schuppert, and Ivo Wallimann-Helmer, eds. 2015. *Social Equality: On What it Means to be Equals*. Oxford: Oxford University Press.
- Fricker, Miranda. 2007. *Epistemic Injustice: Power and the Ethics of Knowing*. New York: Oxford University Press.
- Frost-Arnold, Karen. 2016. "Strategies for Dealing with Online Harassment." Presentation given at the American Philosophical Association Eastern Division Meeting, January 8, Washington, DC. http://www.academia.edu/download/41211817/Frost-Arnold_Strategies_for_Dealing_with_Online_Harassment.pdf
- Frye, Marilyn. 1983. *The Politics of Reality: Essays in Feminist Theory*. Trumansburg, NY: Crossing Press.
- Gelber, Katharine. 2012. "'Speaking Back': The Likely Fate of Hate Speech Policy in the United States and Australia." In *Speech & Harm: Controversies Over Free Speech*, edited by Ishani Maitra and Mary Kate McGowan, 50–71. New York: Oxford University Press.
- Gelber, Katharine, and Luke McNamara. 2016. "Evidencing the Harms of Hate Speech." *Social Identities* 22 (3): 324–341.
- Gray-Donald, David. 2018. "Canada's Right-Wing Rage Machine vs. Nora Loreto." *Briarpatch*, April 15. <https://briarpatchmagazine.com/blog/view/canadas-right-wing-rage-machine-vs-nora-loreto>
- Green, Mitchell. 2016. "Speech Acts: An Annotated Bibliography." *Oxford Bibliographies Online*.
- Hanrahan, Rebecca, and Louise Antony. 2005. "Because I Said So: Towards a Feminist Theory of Authority." *Hypatia* 20 (4): 59–79.
- Harbath, Katie. 2018. "Hard Questions: Social Media and Democracy." *Facebook Newsroom*. January 22. <https://newsroom.fb.com/news/2018/01/hard-questions-democracy/>
- Harris-Perry, Melissa. 2011. "The Epistemology of Race Talk." *The Nation*. September 26. <https://www.thenation.com/article/epistemology-race-talk/>.
- Harwell, Drew. 2019. "Pelosi Videos, Slowed to Make Her Appear Drunk, Spread Across Social Media." *The Washington Post*. May 23. <https://www.washingtonpost.com/technology/2019/05/23/faked-pelosi-videos-slowed-make-her-appear-drunk-spread-across-social-media>.
- Haslanger, Sally, 2012. *Resisting Reality: Social Construction and Social Critique*. New York: Oxford University Press.
- . 2015. "Social Structure, Narrative and Explanation." *Canadian Journal of Philosophy* 45 (1): 1–15.

- Hategan, Elisa. 2004. *Race Traitor: The True Story of Canadian Intelligence Service's Greatest Cover-Up*. Toronto: Incognito Press.
- Herbert, Cassie. 2015. "Precarious Projects: The Performative Structure of Reclamation." *Language Sciences* 52: 131–138.
- . 2018a. "Talking About Slurs." Unpublished Manuscript.
- . 2018b. "Women in Philosophy: Online Misogyny and our Profession." *APA Blog*. July 4. <https://blog.apaonline.org/2018/07/04/women-in-philosophy-online-misogyny-and-our-profession/>
- Herbert, Cassie, and Rebecca Kukla. 2016. "Ingrouping, Outgrouping, and Peripheral Speech." *Journal of the American Philosophical Association* 2 (4): 576–596.
- Hom, Christopher. 2008. "The Semantics of Racial Epithets." *Journal of Philosophy* 105: 416–440.
- . 2012. "A Puzzle About Pejoratives." *Philosophical Studies* 159: 383–405.
- Hominh, Yarran. 2016. "Re-Reading the Declaration of Independence as Perlocutionary Performative." *Res Publica* 22 (4): 423–444.
- hooks, bell. 1994. *Teaching to Transgress: Education as the Practice of Freedom*. New York: Routledge.
- Hornsby, Jennifer. 1994. "Illocution and its Significance." In *Foundations of Speech Act Theory: Philosophical and Linguistic Perspectives*, edited by Savas L. Tsohatzidis, 187–207. London: Routledge.
- Hornsby, Jennifer, and Rae Langton. 1998. "Free Speech and Illocution." *Legal Theory* 4: 21–37.
- Irigaray, Luce. 1985. *Speculum of the Other Woman*. Translated by Gillian C. Gill. Ithaca: Cornell University Press.
- Jacobson, Daniel. 1995. "Freedom of Speech Acts? A Response to Langton." *Philosophy and Public Affairs* 64: 64–79.
- Jane, Emma. 2014. "'Your a Ugly, Whorish, Slut,': Understanding E-bile." *Feminist Media Studies* 14 (4): 531–546.
- Jaschik, Scott. 2016. "The Law Professor Who Answered Back." *Inside Higher Ed*. <https://www.insidehighered.com/news/2016/07/12/law-professor-responds-students-who-complained-about-her-black-lives-matter-shirt>
- Jeong, Sarah. 2018. *The Internet of Garbage*, revised edition. New York: The Verge.
- Jeshion, Robin. 2013. "Expressivism and the Offensiveness of Slurs." *Philosophical Perspectives*. 27: 231–259.
- Kakutani, Michiko. 2018. *The Death of Truth: Notes on Falsehood in the Age of Trump*. New York: Penguin Random House.
- Khan-Cullors, Patrisse, and asha bande. 2018. *When They Call You a Terrorist: A Black Lives Matter Memoir*. New York: St. Martin's Press.
- King Jr., Martin Luther. 1963. *Why We Can't Wait*. New York: Harper & Row.
- Koul, Scaachi. 2017. *One Day We'll All be Dead and None of This Will Matter*. Toronto: Doubleday Canada.
- Kukla, Rebecca. 2014. "Performative Force, Convention, and Discursive Injustice." *Hypatia* 29 (2): 440–457.
- . 2018. "Slurs, Interpellation, and Ideology." *The Southern Journal of Philosophy* 56 (S1): 7–32.
- Kukla, Rebecca, and Mark Lance. 2009. *'Yo!' and 'Lo!': The Pragmatic Topography of the Space of Reasons*. Cambridge: Harvard University Press.
- Kumashiro, Kevin K. 2018. "How Hate Speech Complicates Our Understanding of Bullying." Webinar prepared for the International Bullying Prevention Association. July 24.
- La, Lynn. 2017. "Here's How Trolls Treat the Women of CNET." *CNET*. July 11. <https://www.cnet.com/news/cnet-women-hate-troll-comments/>
- Lance, Mark and Rebecca Kukla. 2013. "Leave the Gun; Take the Cannoli! The Pragmatic Topography of Second-Person Calls." *Ethics* 123: 456–478.
- Langton, Rae. 1993. "Speech Acts and Unspeakable Acts." *Philosophy & Public Affairs* 22 (4): 293–330.
- . 2012. "Beyond Belief: Pragmatics in Hate Speech and Pornography." In *Speech & Harm: Controversies*

- Over Free Speech*, edited by Ishani Maitra and Mary Kate McGowan, 72–93. New York: Oxford University Press.
- . 2017. “Is Pornography like the Law?” In *Beyond Speech: Pornography and Analytic Feminist Philosophy*, edited by Mari Mikkola, 23–38. New York: Oxford University Press.
- . 2018a. “The Authority of Hate Speech.” In *Oxford Studies in Philosophy of Law*, volume 3, edited by John Gardner, Leslie Green, and Brian Leiter, 123–152. New York: Oxford University Press.
- . 2018b. “Blocking as Counter Speech.” In *New Works on Speech Acts*, edited by Daniel W. Harris, Daniel Fogal, and Matt Moss, 144–164. Oxford: Oxford University Press.
- Langton, Rae, Sally Haslanger, and Luvell Anderson. 2012. “Language and Race.” In *The Routledge Companion to Philosophy of Language*, edited by Gillian Russell and Delia Graff Fara, 753–767. New York: Routledge.
- Lapowsky, Issie. 2017. “Trump and the Rise of Digital Hate.” *Wired*. December 2. <https://www.wired.com/story/historians-see-terrifying-echoes-in-trumps-anti-muslim-tweets/>
- Lavender, Paige. 2012. “Rand Paul: Obama’s Views on Marriage Couldn’t ‘Get Any Gay’ (video).” *Huffington Post*. May 12. https://www.huffingtonpost.ca/2012/05/12/rand-paul-obama-gay-marriage_n_1511920.html
- Lawrence, Charles III. 1993. “If He Hollers Let Him Go: Regulating Racist Speech on Campus.” In *Words that Wound: Critical Race Theory, Assaultive Speech and the First Amendment*, edited by Mari Matsuda, Charles Lawrence III, Richard Delgado, and Kimberlé Crenshaw, 53–88. Boulder: Westview Press.
- Lebron, Christopher J. 2017. *The Making of Black Lives Matter: A Brief History of an Idea*. New York: Oxford University Press.
- Lederer, Laura J. 1995. “Pornography and Racist Speech as Hate Propaganda.” In *The Price We Pay: The Case Against Racist Speech, Hate Propaganda, and Pornography*, edited by Laura Lederer and Richard Delgado, 131–140. New York: Hill and Wang.
- Lederer, Laura, and Richard Delgado, eds. 1995. *The Price We Pay: The Case Against Racist Hate Speech, Hate Propaganda, and Pornography*. New York: Hill and Wang.
- Levine, Jon. “The Economist Apologizes for Tweet Asking Whether Transgender People Should Be ‘Sterilized.’” *The Wrap*. March 20. <https://www.thewrap.com/the-economist-apologizes-for-tweet-asking-whether-transgender-people-should-be-sterilized/>
- Levmore, Saul. 2010. “The Internet’s Anonymity Problem.” In *The Offensive Internet: Speech, Privacy and Reputation*, edited by Saul Levmore and Martha C. Nussbaum, 50–67. Cambridge: Harvard University Press.
- Levmore, Saul, and Martha C. Nussbaum, eds. 2010. *The Offensive Internet: Speech, Privacy and Reputation*. Cambridge: Harvard University Press.
- Lewis, David. 1979. “Scorekeeping in a Language Game.” *Journal of Philosophical Logic* 8 (1): 339–359.
- Liebow, Nabina. 2017. “Microaggressions and Moral Responsibility.” Ph.D. diss. Georgetown University.
- Litt, Eden. 2012. “Knock, Knock. Who’s There? The Imagined Audience.” *Journal of Broadcasting & Electronic Media* 56 (3): 330–345.
- Locke, Alain. 1928. “Art or Propaganda?” *Harlem: A Forum of Negro Life* 1 (1): 1–2. Available at: <http://nationalhumanitiescenter.org/pds/maai3/protest/text10/lockeartorpropaganda.pdf>
- Lynch, Michael Patrick. 2016. *The Internet of Us: Knowing More and Understanding Less in the Age of Big Data*. New York: Liveright Publishing Corporation.
- MacKinnon, Catharine. 1987. *Feminism Unmodified: Discourses on Life and Law*. Cambridge: Harvard University Press.
- . 1991. “Pornography as Defamation and Discrimination.” *Boston University Law Review* 71: 793–815.
- . 1993. *Only Words*. Cambridge: Harvard University Press.

- . 1995. "Speech, Equality, and Harm: The Case Against Pornography." In *The Price we Pay: The Case Against Racist Speech, Hate Propaganda, and Pornography*, edited by Laura Lederer and Richard Delgado, 301–314. New York: Hill and Wang.
- Maitra, Ishani. 2009. "Silencing Speech." *Canadian Journal of Philosophy* 39 (2): 309–338.
- . 2012. "Subordinating Speech." In *Speech & Harm: Controversies Over Free Speech*, edited by Ishani Maitra and Mary Kate McGowan, 94–120. New York: Oxford University Press.
- . 2013. "Subordination and Objectification." *Journal of Moral Philosophy* 10 (1): 87–100.
- Maitra, Ishani, and Mary Kate McGowan, eds. 2012. *Speech & Harm: Controversies Over Free Speech*. New York: Oxford University Press.
- Manne, Kate. 2014. "In Ferguson and Beyond, Punishing Humanity." *The New York Times*. October 12. <https://opinionator.blogs.nytimes.com/2014/10/12/in-ferguson-and-beyond-punishing-humanity/>.
- . 2017. *Down Girl: The Logic of Misogyny*. New York: Oxford University Press.
- Matsuda, Mari. 1993. "Public Response to Racist Speech: Considering the Victim's Story." In *Words that Wound: Critical Race Theory, Assaultive Speech and the First Amendment*, edited by Mari Matsuda, Charles Lawrence III, Richard Delgado, and Kimberlé Crenshaw, 17–51. Boulder: Westview Press.
- Matsuda, Mari, Charles Lawrence III, Richard Delgado, and Kimberlé Crenshaw, eds. 1993. *Words that Wound: Critical Race Theory, Assaultive Speech and the First Amendment*. Boulder: Westview Press.
- McGowan, Mary Kate. 2003. "Conversational Exercitives and the Force of Pornography." *Philosophy and Public Affairs* 31: 155–189.
- . 2004. "Conversational Exercitives: Something Else We Do With Our Words." *Linguistics and Philosophy* 27: 93–111.
- . 2009. "Oppressive Speech." *Australasian Journal of Philosophy* 87 (3): 389–407.
- . 2012. "On 'Whites Only' Signs and Racist Hate Speech: Verbal Acts of Racial Discrimination." In *Speech & Harm: Controversies Over Free Speech*, edited by Ishani Maitra and Mary Kate McGowan, 121–47. New York: Oxford University Press.
- . 2018. "On Covert Exercitives: Speech and the Social World." In *New Works on Speech Acts*, edited by Daniel W. Harris, Daniel Fogal, and Matt Moss, 185–201. Oxford: Oxford University Press.
- McKinnon, Rachel. 2018. "The Epistemology of Propaganda." *Philosophy and Phenomenological Research* 96 (2): 483–489.
- Mill, John Stuart. 1859. *On Liberty*. London: John W. Parker and Son, West Strand.
- Mills, Charles. 2005. "Ideal Theory as Ideology." *Hypatia* 20 (3): 165–183.
- Nielsen, Laura Beth. 2012. "Power in Public: Reactions, Responses, and Resistance to Offensive Public Speech." In *Speech & Harm: Controversies Over Free Speech*, edited by Ishani Maitra and Mary Kate McGowan, 148–173. New York: Oxford University Press.
- Nguyen, C Thi. 2018. "Escape the Echo Chamber." *Aeon*. April 9. <https://aeon.co/essays/why-its-as-hard-to-escape-an-echo-chamber-as-it-is-to-flee-a-cult>
- Norlock, Kathryn J. 2017a. "Public Forgiveness; Unforgiving Public." *The Spoke*. May 25. <https://www.wellesley.edu/albright/about/blog/3441-public-forgiveness-unforgiving-public>.
- . 2017b. "Online Shaming." *Social Philosophy Today*. (Online first).
- North, Anna. 2017. "President Trump's Golden Age of Trolling." *New York Times*. July 5. <https://www.nytimes.com/2017/07/05/opinion/trump-cnn-trolls-wrestling-gif.html>
- Nunberg, Geoff. 2016. "The Social Life of Slurs." In *New Work on Speech Acts*, edited by Daniel Harris, Daniel Fogal, and Matt Moss, 237–295. New York: Oxford University Press.
- Nussbaum, Martha C. 2010. "Objectification and Internet Misogyny." In *The Offensive Internet: Speech, Privacy and Reputation*, edited by Saul Levmore and Martha C. Nussbaum, 68–87. Cambridge: Harvard University Press.

- Oxford Living Dictionaries. 2017. "Hate Speech." Accessed 2 November 2017.
https://en.oxforddictionaries.com/definition/hate_speech.
- Paradkar, Shree. 2017. "The Yonge St. Riots of 1992 ... Or Was it an Uprising?" *The Toronto Star*. May 5.
<https://www.thestar.com/news/gta/2017/05/05/the-yonge-street-riot-of-1992-or-was-it-an-uprising-paradkar.html>.
- Phillips, Whitney. 2018. "The Oxygen of Amplification: Better Practices for Reporting on Extremists, Antagonists, and Manipulators." *Data & Society*. May 22. <https://datasociety.net/output/oxygen-of-amplification/>
- Popa-Wyatt, Mihaela, and Jeremy L. Wyatt. 2018. "Slurs, Roles and Power." *Philosophical Studies* 175 (11): 2879–2906.
- Potts, Christopher. 2005. *The Logic of Conventional Implicatures*. Oxford Studies in Theoretical Linguistics. Oxford: Oxford University Press.
- Pratt, Mary Louise. 1986. "Ideology and Speech-Act Theory." *Poetics Today* 7 (1): 59–72.
- Quinn, Zoë. 2017. *Crash Override: How Gamergate (Nearly) Destroyed My Life, and How We Can Win the Fight Against Online Hate*. New York: Public Affairs.
- Ramsey, Franchesca. 2016. "4 Black Lives Matter Myths Debunked." *MTV Decoded*.
https://www.youtube.com/watch?v=iQ_0bqWKO-k
- Ramsey, John Michael. 2013. "How We Do Bad Things with Words: A Multi-level Model of Oppression." PhD diss., University of California, Riverside.
- Rini, Regina. 2017. "Fake News and Partisan Epistemology." *Kennedy Institute of Ethics Journal* 27 (2) supplement: 43–64.
- . 2018a. "How to Take Offense: Responding to Microaggression." *Journal of the American Philosophical Association* 4 (3): 332–351.
- . 2018b. "Social Media and Democracy are Incompatible, Now What?" Paper presented at the Centre for Ethics at the University of Toronto. February 27.
- . 2019. "Deepfakes and the Epistemic Backstop." Unpublished Manuscript (June).
<https://philpapers.org/rec/RINDAT>
- Roose, Kevin. 2018. "'The Daily': The Business of Internet Outrage." *New York Times*. October 31.
<https://www.nytimes.com/2018/10/31/podcasts/the-daily/mad-world-news-facebook-internet-anger.html>
- Ronson, Jon. 2015. *So You've Been Publicly Shamed*. New York: Riverhead Books.
- Saul, Jennifer M. 2017. "Racial Figleaves, the Shifting Boundaries of the Permissible, and the Rise of Trump." *Philosophical Topics* 45 (2): 91–116.
- . 2018a. "Dogwhistles, Political Manipulation, and Philosophy of Language." In *New Work on Speech Acts*, edited by Daniel Harris, Daniel Fogal, and Matt Moss, 360–383. New York: Oxford University Press.
- . 2018b. "Beyond Just Silencing: A Call for Complexity in Discussions of Academic Free Speech." In *Academic Freedom*, edited by Jennifer Lackey, 119–134. New York: Oxford University Press.
- Saslow, Eli. 2018. *Rising Out of Hatred: The Awakening of a Former White Nationalist*. New York: Doubleday.
- Sbisà, Marina. 2002. "Speech Acts in Context." *Language & Communication* 22, 421–436.
- . 2018. "Varieties of Speech Act Norms." In *Dynamics and Varieties of Speech Actions, a theme issue of Poznan Studies in the Philosophy of the Sciences and the Humanities* 112: 23–50.
- Scheffler, Samuel. 2010. *Equality and Tradition: Questions of Value in Moral and Political Theory*. Oxford: Oxford University Press.
- Scheman, Naomi. 1993. "Anger and the Politics of Naming." In *Engenderings: Constructions of Knowledge, Authority, and Privilege*. New York: Routledge.

- Schwartzman, Lisa H. 2002. "Hate Speech, Illocution, and Social Context: A Critique of Judith Butler." *Journal of Social Philosophy* 33 (3): 421–441.
- Sherr, Ian, and Erin Carson. 2017. "GamerGate to Trump: How Video Game Culture Blew Everything Up." *CNET*. November 27. <https://www.cnet.com/news/gamergate-donald-trump-american-nazis-how-video-game-culture-blew-everything-up/>
- Searle, John R. 1976. "A Classification of Illocutionary Acts." *Language in Society* 5 (1): 1–23.
- . 1994. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.
- Sellars, Wilfrid. 1954. "Some Reflections on Language Games." *Philosophy of Science* 21 (3): 204–228.
- Simpson, Robert Mark, and Amia Srinivasa. 2018. "No Platforming." In *Academic Freedom*, edited by Jennifer Lackey, 186–210. New York: Oxford University Press.
- Smith, David Livingstone. 2012. *Less than Human: Why we Demean, Enslave, and Exterminate Others*. New York: St. Martin's Griffin Press.
- . 2017. "Fighting Hate is a Losing Battle." *Boston Globe*. August 29. <https://www.bostonglobe.com/opinion/2017/08/29/smith/i5F9Mf4ZqPohu4oxC6stTP/story.html>
- Solon, Olivia. 2017. "The Future of Fake News: Don't Believe Everything You Read, See or Hear." *The Guardian*. July 26. <https://www.theguardian.com/technology/2017/jul/26/fake-news-obama-video-trump-face2face-doctored-content>
- Stalnaker, Robert. 2002. "Common Ground." *Linguistics and Philosophy* 25: 701–721.
- Stanley, Jason. 2015. *How Propaganda Works*. Princeton: Princeton University Press.
- . 2018. "Reply to Critics." *Philosophy and Phenomenological Research* 96 (2): 497–511.
- Strum, Philippa. 1999. *When the Nazis Came to Skokie: Freedom for Speech We Hate*. Lawrence: The University Press of Kansas.
- Swanson, Eric. [Forthcoming]. "Slurs and Ideology." Unpublished manuscript.
- Táíwò, Olúfẹ̀mí. 2017. "Beware of Schools Bearing Gifts: Miseducation and Trojan Horse Propaganda." *Public Affairs Quarterly* 31 (1): 1–18.
- Talisse, Robert. 2017. "Race and Democratic Virtue with Paul C. Taylor." *Why We Argue*. Humility & Conviction in Public Life. <http://humilityandconviction.uconn.edu/why-we-argue/>.
- Tashman, Brain. 2013. "Robertson: Facebook Should Have a 'Vomit' Button for Pictures of Gay Couples." *Right Wing Watch*. July 8. <http://www.rightwingwatch.org/post/robertson-facebook-should-have-vomit-button-for-pictures-of-gay-couples/>
- Taylor, Keeanga-Yamahtta. 2016. *From #BlackLivesMatter to Black Liberation*. Chicago: Haymarket Books.
- Technau, Björn. 2018. "Going Beyond Hate Speech: The Pragmatics of Ethnic Slur Terms." *Lodz Papers in Pragmatics* 14 (1): 25–43.
- Thomason, Richmond H. 1990. "Accommodation, Meaning, and Implicature: Interdisciplinary Foundations for Pragmatics." In *Intentions in Communication*, edited by Philip R. Cohen, Jerry Morgan, and Marta E. Pollack, 325–364. Cambridge: MIT Press.
- Thompson, A.C., Ali Winston, and Darwin BondGraham. 2017. "Racist, Violent, Unpunished: A White Hate Group's Campaign of Menace." *ProPublica*. October 19. <https://www.propublica.org/article/white-hate-group-campaign-of-menace-rise-above-movement>
- Tirrell, Lynne. 1993. "Definition and Power: Toward Authority without Privilege." *Hypatia* 8 (4): 1–34.
- . 1997. "Language and Power." In *A Companion to Feminist Philosophy*, edited by Alison M. Jaggar, and Iris Marion Young, 139–152. Malden: Blackwell.
- . 1999. "Derogatory Terms: Racism, Sexism, and the Inferential Role Theory of Meaning." In *Language and Liberation: Feminism, Philosophy, and Language*, edited by Christina Hendricks and Kelly Oliver, 41–79. Albany: SUNY Press.
- . 2012. "Genocidal Language Games" In *Speech and Harm: Controversies over Free Speech*, edited by Ishani Maitra, and Mary Kate McGowan, 174–221. Oxford: Oxford University Press.

- . 2017. "Toxic Speech: Towards an Epidemiology of Discursive Harm." *Philosophical Topics* 45 (2): 139–161.
- . 2018a. "Toxic Speech: Inoculations and Antidotes." *Southern Journal of Philosophy* 56 (S1): 116–144.
- . 2018b. "Authority and Gender: Flipping the F-Switch." *Feminist Philosophy Quarterly* 4 (3).
- Tufekci, Zeynep. 2017. *Twitter and Teargas: The Power and Fragility of Networked Protest*. New Haven: Yale University Press.
- Valenti, Jessica. 2016. *Sex Object*. Dey Street Books.
- Veilleux, Fred. 1995. "Indians are a People, Not Mascots." In *The Price we Pay: The Case Against Racist Speech, Hate Propaganda, and Pornography*, edited by Laura Lederer and Richard Delgado, 45–54. New York: Hill and Wang.
- Waldron, Jeremy. 2012. *The Harm in Hate Speech*. Cambridge: Harvard University Press.
- Wartenberg, Tom. 1988. "The Situated Conception of Social Power." *Social Theory and Practice* 14 (3): 317–343.
- Warzel, Charlie. 2019. "Mass Shootings Have Become a Sickening Meme." *New York Times*, April 28. <https://www.nytimes.com/2019/04/28/opinion/poway-synagogue-shooting-meme.html>
- Washington Post staff. 2015. "Full text: Donald Trump Announces a Presidential Bid." *Washington Post*. June 16. https://www.washingtonpost.com/news/post-politics/wp/2015/06/16/full-text-donald-trump-announces-a-presidential-bid/?utm_term=.e8691f46435b.
- West, Caroline. 2012. "Words that Silence? Freedom of Expression and Racist Hate Speech." In *Speech and Harm: Controversies over Free Speech*, edited by Ishani Maitra and Mary Kate McGowan, 222–248. Oxford: Oxford University Press.
- West, Lindy. 2016. *Shrill: Notes from a Loud Woman*. New York: Hatchette Books.
- Whiting, Daniel. 2013. "It's Not What You Said, It's the Way You Said It: Slurs and Conventional Implicatures." *Analytic Philosophy* 54 (3): 364–377.
- Wolff, Jonathan. 1998. "Fairness, Respect, and the Egalitarian Ethos." *Philosophy and Public Affairs* 27: 97–122.
- . 2016. "How Propaganda Works by Jason Stanley." *Analysis* 76 (4): 558–560.
- . 2019. "Equality and Hierarchy." *Proceedings of the Aristotelian Society* 119 (1): 1–23.
- Women's Media Center Speech Project. www.womensmediacenter.com/speech-project/research-statistics/
- Wood, Graeme. 2017. "How Richard Spencer Became an Icon for White Supremacists." *The Atlantic*. <https://www.theatlantic.com/magazine/archive/2017/06/his-kampf/524505/>
- Young, Iris Marion. 1997. "Asymmetrical Reciprocity: On Moral Respect, Wonder, and Enlarged Thought." *Constellations* 3 (3): 340–363.
- . 2004. "Responsibility and Global Labor Justice." *The Journal of Political Philosophy* 12 (4): 365–388.
- . 2006. "Responsibility and Global Justice: A Social Connection Model," *Social Philosophy & Policy* 23: 102–130.
- . 2011. *Justice and the Politics of Difference*. Princeton: Princeton University Press.
- . 2013. *Responsibility for Justice*. Oxford: Oxford University Press.
- Zheng, Robin. 2018. "Bias, Structure, and Injustice: A Reply to Haslanger." *Feminist Philosophy Quarterly* 4 (1). Article 4.