When Does Speech Perform Regulable Action?

Weston, Daniel

International Journal of Language and Law

Published: 01/11/2022

Publisher's PDF, also known as Version of record

International Journal of
Language & Law

Weston, JLL 11 (2022): 78–97
www.languageandlaw.eu

When Does Speech Perform Regulable Action?

— A Critique of Speech Act Theory's Application to Free Speech Regulation

*Daniel Weston**

Abstract
This paper examines the application of speech act theory to free speech regulation and criticises the idea that an understanding of speech which performs speech acts can be of use in identifying regulable speech. It traces the legal application of speech act theory from its initial uses to the more contemporary. In response, this paper seeks to demonstrate that this influential legal application reaches conclusions against the core insight of speech act theory – that *all* speech performs actions in the relevant, illocutionary and performative, sense. Consequently, an arbitrary method in regulating speech has taken firm hold in contemporary free speech theory, through which some speech is erroneously perceived to be more like a form of speech act than speech proper. I examine the lessons of speech act theory alongside this free speech literature to conclude that we should not ask *whether* an utterance is an act but instead *what kind* of act it is, with the goal of refocusing on normative questions pertaining to speech regulation.

Keywords
free speech, speech act theory, illocutionary acts, performativity, hate speech, pornography

Submitted: 12 October 2021, accepted: 1 April 2022, published online: 1 November 2022

* *Daniel Weston*: University of Law, daniel.weston@law.ac.uk or dan.weston1990@gmail.com. I am very grateful for the attentive reading and irreplaceable feedback on drafts/the ideas of this work from Stefano Bertea, Peter Cumper, Stephen Pethick, Claiton Fyock, Stephen Riley and the anonymous reviewers for this paper, as well as everyone I discussed these ideas with at University of Leicester during my PhD (Weston, 2020), upon which this article is based, and which represents a distilled version of that work. I also appreciate the opportunity to present and receive feedback on an early draft of this work at the Society of Legal Scholars Conference 2019.

1. Introduction

One question at the core of a principle of free speech is what counts as 'speech' for legal protection and what is instead more akin to unprotected 'action' or 'conduct'. This conceptual distinction has attracted considerable legal interest in attempts to gain a stronger understanding of the appropriate regulation of speech. Frederick Schauer, for instance, has asked "whether the kind of distinction between speech and action that is necessary to any principle of free speech can in fact be sustained" (2015: 430). He suggests that "we must subject to critical analysis just what it means to draw a distinction between speech and action" (2015: 430). The literature on the proper demarcation between speech and action for legal regulatory purposes is extensive and varied, with Amy Adler remarking that "the problem of dividing speech from action consumes a significant part of case law and scholarship" (2001: 973).

This paper hones in on one approach that free speech theorists have employed to draw the distinction: through the application of speech act theory – a philosophy of language often summarised by the phrase that "to *say* something is to *do* something" (Austin, 1962: 12). Herein I will criticise various claims that some speech might be legally regulated on the basis it performs a kind of speech action as "performative" (Austin, 1962: 4) or "illocutionary" (Austin, 1962: 98) utterances – terms I shall elucidate herein. These claims are experiencing a renaissance in the context of hate speech and pornography regulation (which are often considered hand-in-hand in speech act applications), with particularly thorough analyses made by Mary Kate McGowan (2012; 2019), Rae Langton (2009, 2012) and Ishani Maitra (2012). Such a speech act approach to regulating speech is also found in relevant foundational work undertaken by Kent Greenawalt (1989) and Catharine MacKinnon (1993).

Though existing legal applications of speech act theory offer distinct points, this paper connects them to a shared underlying commitment to the regulation of *speech which performs acts*. Greenawalt makes an early and highly influential application of speech act theory to free speech (1989: 57). John Wirenius for instance notes that the speech and action distinction in free speech theory "has generally not been explored in a fruitful manner. An exception is Kent Greenawalt" (2004: 147). Greenawalt distinguishes between speech which involves "ways of doing things, not of asserting things" (1989: 58). In applying this distinction, Greenawalt draws on speech act theory, noting that "[m]odern recognition of the importance of this use of language owes a great deal to J. L. Austin, who coined the term 'performance utterances'" (1989: 58).

A significant body of later work employs speech act theory with specific speech targets in mind: the regulation of hate speech and pornography (these are both broad terms, but as will become apparent in Section Four, my critique applies across varied constructions). McGowan argues for instance that "some racist hate speech ought to be regulated because it constitutes an (otherwise illegal) act of racial discrimination" (2012:

122). Similarly, Langton asks of speech act theory: "[c]an these stories told by philosophers shed light on hate speech?" (2012: 73). She answers "probably, yes" (2012: 73) and accordingly explores the legal implications that pornography and hate speech are a "kind of speech act" (2009: 25). Along the same line of argument Maitra writes that "the very act of producing speech of that kind [hate speech] just *is* a subordinating (speech) act" (2012: 98). These arguments regarding pornography and hate speech are a continuation of work done by MacKinnon (Maitra 2012: 94; Langton 2009: 2), who argues for a renewed understanding of regulating such speech on the basis they operate "more in active than in passive terms, as constructing and performative rather than as merely referential or connotative" (1993: 21).

Though it may seem like a surprising charge to levy at the outset, and it is a claim that requires added nuance, this paper argues that these theorists fundamentally misunderstand the implications of the fact that *all* speech performs (speech) actions. I shall aim to show it is a misguided question to ask *whether* speech is "doing something" (Austin, 1962: 25), and emphasise instead that free speech theory should focus on *what* it is doing. I shall argue that normative questions – that is, legal, moral or political arguments regarding which speech ought to be regulated – cannot be assisted, mitigated or avoided by this particular speech act categorisation, at least as currently postulated. As this paper aims to show, this misconception runs through some aspects of the earliest free speech applications of speech act theory to the very recent. The error is more deeply embedded in contemporary work in hate speech than in its original application, but a consistent misapplication runs through. Having shown this, this paper then seeks to provide reason to resist the idea that the conflict between free speech and speech regulation can be solved by these particular appeals to speech act theory.

Section Two introduces the fundamentals of speech act theory insofar as it is needed to explain my critique and the claims made by my interlocutors. This Section is structured to provide a legal reader without any familiarity with speech act theory the required knowledge to traverse my interlocutors' claims and my rebuttal. For those with prior familiarity with speech act theory, it might be taken as an attempt to defend my interpretation and demonstrate its widespread support in the mainstream philosophical literature. Those who need no context or persuasion on the idea that all speech involves illocutionary acts might reasonably begin with Section Three. Section Three picks out one erroneous aspect of Greenawalt's theory that has blossomed in some contemporary free speech literature. Readers most interested in the regulatory implications for hate speech and pornography might begin directly with Section Four, which examines these particularly, but Section Three shows the error I discuss is not an isolated feature of some hate speech/pornography theory, but is in fact to some degree embedded in the earliest applications of speech act theory to speech regulation. In a sense, the progression of this paper can be understood as a chronological analysis from the relevant aspects of speech act theory in philosophy, into its first uses in free speech theory, and proceeding up to its contemporary misapplications.

2. Speech Act Theory, Performativity and Illocutionary Acts

Since this paper draws on some technical terms from speech act theory, and my contention hinges on an understanding of it, it is necessary to briefly set out its relevant ideas at the outset. While speech act theory has a great depth to it, one need not venture too far out to establish the relevant principles for my response to its legal employment. Austin conceptualised and invented speech act theory (Cerf, 1969: 351–352) through which he challenged a preoccupation with "the assumption of philosophers that the business of a 'statement' can only be to 'describe' some state of affairs, or to 'state some fact', which it must do either truly or falsely" (Austin, 1962: 1). His considerations took place in the context of a particular propensity within philosophy of language (at the time) to focus mainly on the truth or falsity of statements, and a corresponding neglect of the way speech is used to *do* things in the world (Austin, 1962: 6).

In response to this philosophical state of affairs Austin initially proposed a difference between what he called a constative and a performative, the former being "a statement" (1962: 6) and the latter "doing something" (1962: 25). The dichotomy can be explained by the idea that some speech conveys things and some speech enacts changes in the world by achieving/doing things. The difference is illustrated by the basic example of a statement/constative such as "coffee increases productivity" as contrasted with a performative "get me a double espresso". The former is an archetypical truth-assessable statement, and the latter an order or request (or something else, dependent on context). Alternatively, one can consider the constative "the demonisation of dietary fat has contributed to increased obesity" and the performative "never darken my doorstep again!" On the face of it, the former makes a truth-conditional claim, and the latter tries to affect some change in the world or do something (e.g. banish someone from one's life).

To make a preliminary legal remark, this initial distinction between constatives and performatives enables the categorisation of some speech as especially conduct-like, and one might rely on this classification to regulate performatives without impinging (meaningfully) on a principle of free speech. If it is possible to remove some forms of speech from the free speech equation by distinguishing between speech which makes claims (constatives) and speech which does things (performatives), it would be a powerful conceptual tool for deciding what to regulate. For example, one could distinguish inciting a crowd to attack someone on the street by pointing at them and shouting "Get that traitor!" versus a political argument such as "Donald Trump was the best President the USA has ever had". The former would be, within such a framework, a performative which does something, and the latter would be able to be categorised as a constative which says something, which could assist in solving regulatory questions.

Though Austin initially endorsed this distinction between constatives – utterances which make claims about the world – and performatives – utterances which achieve acts in the world – he later abandoned it. Rather than some speech being a truth-oriented

constative and some the performance of an action he considered that any given utterance is both. He writes that "stating something is performing an act just as much as giving an order or giving a warning" and that "when we give an order or a warning or a piece of advice, there is a question about how this is related to fact" (1979: 251). This leads him to conclude that "in its original form our distinction between the performative and the statement is considerably weakened, and indeed breaks down" (1979: 251). It is this observation that led him to develop speech act theory more fully in place of a simpler notion of performativity.

It is therefore necessary to distinguish Austin's original theory of *performativity* from his theory of *speech acts*. Speech act theory instead maintains that all speech has three aspects to it. The idea is summarised as identifying that any utterance a) transmits information b) performs actions and c) has consequential effects. These three aspects represent simple technical notions: the locutionary aspect of speech reflects an utterance's factual aspects, the illocutionary aspect represents the action(s) conducted by an utterance, and the perlocutionary represents the effects of a given utterance (Austin, 1962: 100–101). All utterances therefore convey something, do something and cause something (or indeed many things). Austin accordingly writes that "[t]o perform a locutionary act is in general, we may say, also and eo ipso to perform an illocutionary act, as I propose to call it" (1962: 98).

Austin's conclusions in this respect are shared throughout speech act theory beyond his original treatment. In support, Keith Graham writes that Austin "is therefore led to develop the theory of illocution as a *replacement* for the first theory" (Graham 1977: 54). Similarly, Walter Cerf concludes that "as soon as Austin has established the apparently clear-cut distinction between constatives and performatives, he proceeds to show that the distinction breaks down" (Cerf, 1969: 353). Elena Collavin provides contemporary support: "at the core of speech act theory […] all utterances amount to the execution of an act" (2011: 373). She notes of Austin that "[a]fter having first created the constative/performative dichotomy, he ultimately erodes it and argues that all utterances are in fact used to perform speech acts" (2011: 377). Even speech act theorists who are overall critical of Austin's conclusions, such as David Holdcroft, acknowledge the fundamental point that "one cannot say something without performing some illocutionary act or other" (1978: 15).

Among these developments to speech act theory, John Searle is the most significant post-Austin authority of speech act theory, with Collavin noting "Searle's systemization and development of Austin's ideas has been very influential, to the point that Searle's interpretation of the theory is at times taken as the definitive view of speech acts" (2011: 377). Despite technical refinements/contestations between his work and Austin's, this fundamental observation remains consistent throughout his and Austin's work, with Searle writing that "all linguistic communication involves linguistic acts" (1969: 16) and that "communication necessarily involves speech acts" (1969: 17). The breadth of this claim is clear when he writes:

> Some of the English verbs and verb phrases associated with illocutionary acts are: state, assert, describe, warn, remark, comment, command, order, request, criticize, apologize, censure, approve, welcome, promise, express approval, and express regret. (1965: 221)

This list by Searle illustrates (though not exhaustively) the wide variety of things one may do with words. This can be initially contextualised within the previous basic examples offered in this paper. Riling up a crowd with the supposed performative "Get that traitor!" involves factual claims (that the target is a traitor, that they deserve to be got, among other things) and the constative "Donald Trump was the best President" involves (depending on the widely possible contexts) a variety of possible (speech) actions: persuading, joking, intimidating, concluding, asserting, censuring, etc. That only *some* utterances do things/perform actions then is called into question by speech act theory. Instead, all utterances perform, or try to perform (a failed speech act is described by Austin as "infelicitous" (1962: 16)), a great variety of different actions. I develop the legal dimensions of this in subsequent sections.

Though the theorists this paper critiques acknowledge that Austin and successors reject performativity theory, they nonetheless pursue theories of speech regulation on the basis such speech performs speech actions – as either illocutionary acts or performatives. Instead, this paper argues that since all speech *acts* (in the relevant way: illocutionarily, or as a performative), legal regulability cannot be resolved by arguing that some speech performs speech actions. In the limited sense meant by speech performing 'actions' some speech may perform actions which warrant regulation. However, such questions of speech regulation must be tackled with specific normative reasons, as opposed to through categorical reliance on this faulty distinction. Given the demonstrable veracity with which speech act theory rejects the distinction between speech which acts and that which does not, this is a likely surprising accusation of neglect to make of free speech theory, the demonstration of which I now turn.

3. Situation-Altering Utterances and Performativity

An early argument for regulating speech on the basis that it is better understood as a kind of speech action can be found in Greenawalt's theory of situation-altering utterances (1989). He is not concerned with any particular kind of speech, (such as hate speech or pornography, explored in Section Four), but instead a broad taxonomy of general "utterances […] outside the scope of a principle of free speech" (1989: 58). The basis of this taxonomy is that "[s]uch utterances are ways of doing things, not of asserting things, and they are generally subject to regulation on the same bases as most noncommunicative behaviour" (1989: 58). In order to explain this, he draws on the philosophy of language, writing that "[a]s is conveyed by the slogan of philosophers of language, 'the dif-

ferent uses of language,' the point of many communications is neither to transmit information nor to assert values" (1989: 57). While there is a legitimate divergence from speech act theory in Greenawalt's work when it comes to *assertive* speech – which is an aspect of his argument I develop elsewhere (Weston, 2022) – in this Section I argue his attempt to employ a modified notion of performativity to inform speech regulation does not evade the problems that led speech act theory to abandon it in the first place (namely that it fails to create a distinction). While Greenawalt's theory is not holistically disputed by the clarification I offer here, it will become apparent how the error in question has proliferated since his writing on the matter and understanding this begins with his work.

In order to refine performativity for legal employment he introduces the notion of situation-altering utterances: "[t]he central idea about situation-altering utterances […] is that they actually change the social world in which we live" (1989: 58). He contrasts situation-altering utterances with "statements of fact and value" (1989: 239) which, he argues, are more relevant for free speech protection because "[a] free speech principle most obviously covers factual propositions of a general sort, those asserted in statements" (1989: 43). He relates this dichotomy (between situation-altering utterances which do things in the world and statements of fact and value) to speech act theory, writing that "[m]odern recognition of the importance of this use of language owes a great deal to J. L. Austin, who coined the term 'performance utterances'" (1989: 58). A core idea for his work is therefore that a free speech principle is and ought to be more prepared to regulate instances of speech which "actually change the world" (1989: 58) rather than "factual propositions" (1989: 43).

As will be apparent, aspects of Greenawalt's claim requires careful distinction from the scope of my critique. He further argues that:

> Forms of communication that are not primarily designed to express facts, values, and feelings, such as words of agreement, promises, offers, and instructions, raise vital issues about the boundaries of speech. (Greenawalt, 2015: 208)

It is true that such utterances may be doing a *particular* thing, but this cannot be theorised toward a categorical investigation into whether they are primarily "ways of doing things" (1989: 58) *generally*. While it is the case that some of the things speech is capable of doing through what it communicates – such as particular offers, agreements or promises – may (in some contexts) be justifiably regulable, this cannot be linked to whether they say or do things, only that what they are specifically doing is regulable. If I 'agree' to pay a gardener if they maintain my shrubbery, I am communicating to them that I will do so, that I have the money, that it will be in a timely fashion, and in so doing I am performing an illocutionary act of agreeing. Alternatively, if I 'instruct' a terrorist how to create a bomb out of household items, I am communicating the way in which this is done, and performing – in a very broad sense – an act of instruction. (These examples might be reversed, I might 'agree' to sell bomb parts to a terrorist, or 'instruct' a gardener to do a

certain thing to a certain tree.) Splitting up these utterances into whether they *change the world by doing things*, or instead state things inherits the lack of distinction encountered by speech act theory in defining performativity.

Greenawalt acknowledges that a direct legal transplant of the performative/constative distinction would be inappropriate:

> My thesis about the nonapplication of the principle of free speech might be regarded as a thesis about utterances that are performative, but good reasons support choosing the term 'situation-altering' instead. (1989: 58)

Accordingly, situation-altering utterances are a modification to Austin's theory of performativity as opposed to a direct importation. As Greenawalt says:

> [Austin] grew increasingly sceptical of any sharp distinction between performative utterances and "constative" utterances […] he may ultimately have doubted that there remained a distinctive class of performative utterances. (1989: 58)

As a result of this, he argues that "[t]he category of utterances I discuss is, accordingly, much narrower than Austin's category of performatives" (1989: 58). Therefore, Greenawalt is not directly relying on performativity but borrows from it in creating his own distinction between situation-altering utterances and statements of fact and value, with which he aims to justify the regulation of the former.

In some respects, this diversion from performativity succeeds, since the notion of situation-altering utterances in part concerns itself with speech that is particularly important to the creation of obligations – in the sense that some speech imposes stronger obligations on participants to act as a result of what was uttered. As Greenawalt writes, "[t]he conventions of language and of ordinary social morality make certain utterances, such as promises, count as far as one's moral obligations are concerned" (1989: 58). In illustration, with respect to agreements, he observes that:

> When two people have agreed to do something, each has undertaken an obligation toward the other to perform the task, an obligation that did not exist before the agreement was made. (1989: 63)

It stands to reason that some speech will involve the generation of stronger obligations on oneself or an interlocutor than other speech. If one were to remark to their partner "the toaster is broken", such a remark, in the right context, imposes some sort of minor obligation on either person to replace it when they are next at the shops. In contrast, if an armed mugger says "give me your money", their utterance is clearly particularly impelling on the recipient. Greenawalt is therefore surely correct to say that some speech will impose greater obligations (on another or oneself) to act than other speech. This however is an entirely separate question as to whether the speech is more or less the doing of actions; instead, it is to observe that *some* of the actions done by speech 'do' a specific thing – impose strong obligations – and that the law ought to be concerned with these kinds of speech actions (though I do not engage with that question further here).

This is therefore not to say that some of the categories Greenawalt highlights as problematic for a principle of free speech such as certain kinds of threats (1989: 249), offers

(1989: 241), or agreements (1989: 79) are not justly regulated, but rather that this regulation cannot be justified by reference to the fact it performs speech acts or does things. It is noteworthy that Greenawalt shares a normative commitment to speech regulation (1989: 40) though my critique differs with respect to what that entails. Offers, agreements and threats are simply the kind of performative the law might be (in some circumstances) normatively concerned with and willing to regulate. For example, it is straightforward to understand the reasons behind restricting speech offering the sale of military weapons to civilians. We might therefore ask what specifically is being *offered* ("want to buy X illegal goods?"), *agreed* ("I'll forge his signature") or *threatened* ("your money or your life!"), and say in those instances the speech involved 'does' something that is harmful enough to regulate. This can hardly be attributed to it doing *something*, which can perhaps more readily be seen as problematic when contrasted with ordinary acts. If we momentarily put the notion of speech acts to the side and consider ordinary actions we would not, for instance, be tempted to say that robbing a bank was *more* of an action than going to work. A bank heist is criminally regulated because of *what* one does when one does it, not because it is more of an action than going to work. If it is doing more, this can only be in the sense of its effects being more significant within a normative framework, rather than because it is more of an action independently of any such analysis. This is very clear with respect to ordinary action; the opposite belief with respect to regulating speech acts has taken hold due to misunderstanding how all utterances involve locutionary and illocutionary acts, to which I now turn.

A lingering concern at this stage may be whether what I critique here is overly linguistic or pedantic. After all, if Greenawalt's theory has merit to it when clarified as I suggest, what does it matter if this mistake occurs? The next Section will show, in an important and contemporary context, what is at stake with regard to the clarification I offer. A spiritual continuation, but distinct form, of Greenawalt's argument has been taken up in the context of hate speech and pornography regulation with particular vigour. Though Greenawalt's theory is a general one not only concerned with explaining particular kinds of speech regulation, his reliance on the significance of speech "doing things" (1989: 58) is a common core between his and more recent theories.

4. Hate Speech and Pornography as Performative and Illocution

While Greenawalt's theory seeks to establish a general taxonomy of different kinds of speech to aid legal regulation, there is a similar but distinct body of literature targeting pornography and hate speech as particularly ripe examples of the ways speech is used to do things. An important early advocate of thinking of pornography and hate speech in

terms of the conduct it performs is MacKinnon. Judith Butler is another notable early advocate for understanding speech via performativity (1997) but for the focused purpose of my argument here, the particulars of her position are distinct to MacKinnon's. More recent applications of speech act theory to free speech also acknowledge MacKinnon's influence on developing a speech act theoretical account of speech regulation (Maitra, 2012: 94; Langton, 2009: 2). In turn, MacKinnon writes in reference to some of this new work that it develops her argument of what "pornography does into the simpler terms that analytic philosophers of language find less possible to reject and deny out of hand" (2012: viii-ix). While it is not my objective to deny these arguments out of hand, I propose that there is good reason to be sceptical of the legal argument that some speech is better understood in terms of what it does as a speech action.

In the context of pornography other legal theorists have taken a critical position on some of these uses of speech act theory, such as Adler, who defends a more firm doctrinal distinction between speech and action, writing that "developments in child pornography law have subverted traditional First Amendment principles that separate speech from conduct" (2001: 926). Alon Harel takes a different line of critique by arguing that it is generally an unimportant distinction: "the speech/act distinction has very little relevance, if any, to the justifiability of regulating speech" (2011: 14). The specifics of these arguments are distinct with radically different conclusions and not explored further here as I do not deal with the speech/action distinction as a whole, nor does my critique necessitate a reinforcement or abandonment of it more broadly.

MacKinnon takes an interest in the performative nature of both pornography and hate speech, remarking of pornography – which in her analysis is a form of hate speech (MacKinnon, 1993: 18; Langton, 2012: 78–79) – that it "requires understanding it more in active than in passive terms, as constructing and performative rather than as merely referential or connotative" (1993: 21). This is similar in spirit to the analysis undertaken by Greenawalt in his articulation of situation-altering utterances, with a more specific application. My broad response is therefore similar: pornography might be performative *in a way that justifies* its regulation, but this is a distinct question, and not informed by virtue of being a performative speech action, but rather because of the kind of performative speech action it is. MacKinnon does not limit her considerations to pornography alone, using speech act theory in application to speech that relates to inequality, and lists a variety of examples of when words are better understood as the performance of acts:

> Words unproblematically treated as acts in the inequality context include "you're fired," "help wanted – male," "sleep with me and I'll give you an A," "fuck me or you're fired," "walk more femininely, talk more femininely, dress more femininely, wear makeup, have your hair styled, and wear jewelery," and "it was essential that the understudy to my Administrative Assistant be a man." These statements are discriminatory acts and are legally seen as such. (MacKinnon, 1993: 13–14)

Her identification of these statements as discriminatory acts in order to argue for their regulation is contingent on it being meaningful that they "can constitute actionable discriminatory acts" (MacKinnon, 1993: 14). This rather gets things the wrong way around as whether we call them discriminatory acts or discriminatory statements is ultimately trivial. If they are discriminatory, and discriminatory *in a way* that warrants regulation, this is an entirely different – normative – matter; not enhanced by thinking of them as acts in a speech act sense. Her pursuit of this re-categorisation culminates in her claim that "[a]t stake in constructing pornography as 'speech' is gaining constitutional protection for doing what pornography does: subordinating women through sex" (MacKinnon, 1993: 29). The interesting question under this query ought to be whether pornography does subordinate women, and if so, in a way that warrants legal intervention. MacKinnon may or may not be right that pornography does subordinate women, this is a matter specific to the effects of pornography, but when considering its regulation, no weight can be accorded to the fact that it has performative, as well as communicative, aspects. We ought to resist the claim that there should be anything at stake (in speech act terms) in "constructing pornography" (MacKinnon, 1993: 29) as either speech or an act.

MacKinnon extends her considerations to a variety of utterances, taking a wide definition of hate speech, and offers a paradigm instance of her point: "[i]f ever words have been understood as acts, it has been when they are sexual harassment" (1993: 45). In terms of my rebuttal, it matters little which specific example is considered, since in each case she is importing her disfavour of a specific performative onto performatives generally. From a legal regulatory perspective, the only coherently interesting claim can be that sexual harassment acts in a certain way that justifies regulation. To impute these observations to the fact that they are performative *generally*, and promote a theory of regulation which regulates performative speech, misunderstands the core lesson of speech act theory and threatens to shift speech illegitimately outside legal protection.

The normative arguments for or against regulating a kind of speech will be distinct to each instance. To use MacKinnon's examples from the above, there are potential normative reasons to regulate various forms of harassing speech, pornography and discriminatory/hate speech which would be a matter of specific argument, with a multitude of complexities within those categories. To illustrate the point via hate speech, as put by Caleb Yong, a uniform policy in response to hate speech is insufficiently precise and therefore "[t]he appropriate response to different kinds of hate speech will differ from case to case" (2011: 386). This results (quite properly) in a number of arguments for and against the regulation of certain kinds of speech even within one broad category such as hate speech.

There is also complexity and room for substantial disagreement regarding available regulatory responses to any such speech, even if what counts as hate speech can be agreed upon. For instance, Katharine Gelber has:

> [D]eveloped the idea of a policy of speaking back, in which individuals who are the targets of hate speech are provided with the institutional, educational, and material support to enable them to speak back[.] (2012: 51)

Similarly, Ioanna Tourkochoriti, focusing on different potential regulatory responses, "distinguishes between appropriate and non-appropriate avenues the state could pursue in order to eliminate prejudice" (2020: 34). As a result, myriad contextual arguments are made and required about all kinds of potential hate speech and available responses, ranging from the most subtle forms of linguistic discrimination to the more overt. These remarks do not even account for other forms of speech such as pornography or those explored in Section Three such as threats or agreements which all come with their own normative specificity. My intent therefore in this paper is not to wade into the merits of any such arguments in relation to any of the speech phenomena considered herein, but simply to reinforce that these are the important questions *in contrast* to a focus on performative speech.

It is important to clarify, in the interests of avoiding a straw man of her position, that MacKinnon acknowledges that she is "not saying that pornography is conduct and therefore not speech, or that it does things and therefore says nothing and is without meaning" (1993: 30). The claim she makes is more specific and complex than that – not that speech magically becomes conduct by invoking speech act theory – but that some speech is particularly performative or illocutionary, even though it is still 'speech' in the ordinary sense of the word. Her claim is not as trivial as the idea that pornography or hate speech (in the various forms she considers it) is not, literally, speech. Similarly, McGowan acknowledges that the hate speech she identifies (to which I turn momentarily) is still "speech in the ordinary sense" (2012: 125), and Langton also that "[t]o say that pornography is a kind of act is not to say that pornography is conduct, and nothing that I say will turn on that claim" (2009: 28). This is important to acknowledge for all the work I analyse herein, since my critique is not contingent on such a false construction.

Despite invoking linguistic theory to argue that some speech is relevantly performative for regulatory purposes, MacKinnon has in general terms recently written of her older work that "[t]he reality of pornography, not its abstract or linguistic features, was the ground of the confrontation with the buried harms of 'speech'" (2012: vi). Though I disagree with her particular application of performativity theory, there is reason to think that my conclusion is supportive of her general thrust here: that we should pay less attention to linguistic categorisation and more to the reasons for or against regulating/protecting speech. Given that linguistic categorisation is used to (I argue, arbitrarily) isolate and regulate speech – in place of analysis of its harms – it would be my hope that my conceptual adjustments would move the discussion back to a normative approach to regulating speech. As it stands though, the claims I critique currently do pursue a framework of regulating via linguistic features on the basis that this kind of speech is performative and/or illocutionary, and that this enhances arguments for its regulation.

MacKinnon's fundamental interest in performative and illocutionary speech runs through contemporary arguments regarding the regulation of pornography and hate speech. I now turn to discuss three recent theorists who continue to argue that hate speech and pornography are regulable due to their performance of illocutionary acts. Each takes the ideas slightly differently, but my concern here critiques them insofar as they make speech act theory a significant premise of their identification of hate speech/pornography as especially illocutionary or performative. In reference to these successors to her work, MacKinnon writes "[a]n intrepid and insightful group of analytic philosophers has made a real contribution to thinking on the subject of speech and its harms" (2012: viii).

Langton is one of these advocates, and takes up this argument writing that "MacKinnon […] wants to attend not simply to the content of pornographic speech, not simply to its effects, but to the actions constituted by it" (2009: 26–27). She defends MacKinnon's claim that pornography is an "act against women" (Langton, 2009: 4; 1993: 11) and that it is therefore "a speech act, if pornography is [also] speech" (2009: 4). Accordingly, Langton attempts to synthesise the speech aspect of pornography with the actions it performs in remarking that:

> Pornography is speech […] Pornography is a kind of act […] Put these together and we have: pornography is a kind of speech act. In what follows I take this suggestion seriously. (2009: 25)

She aims to understand the regulation of pornography better by considering it as a speech act, and says that "once we consider pornographic images and texts as speech acts, we are in a position to apply to them Austin's distinction between locutionary, illocutionary and perlocutionary acts" (2009: 30). Consequently, Langton focuses on the illocutionary acts that pornography may perform, and which may therefore constitute an act of subordination, rather than merely being speech that causes subordination: "pornography can have the illocutionary force of subordination, and not simply have subordination as its locutionary content, or as its perlocutionary effect: *in* depicting subordination, pornographers subordinate" (Langton, 2009: 34). Linking this to the underlying philosophy of language and speech acts (2009: 27), she harnesses Austin to focus legal attention on the "actions constituted" (2009: 27) by words – in this case, an act of subordination (2009: 34). As a result, she argues that "[b]esides depicting and causing subordination […] pornography *is*, in and of itself, a form of subordination" (2009: 26). Rather than pornography *resulting in* the subordination of women, she argues that it "determines women's inferior social status" (2009: 29) as an illocutionary act, asking, (and answering): "[c]an a speech act be an illocutionary act of subordination? The answer, I think, is yes" (2009: 34).

This claim is not as controversial as she sets out, however, since the answer is of course it can be an illocutionary act, but the fact that it *is* is trivially true. Speech can be used to perform, as considered in Section Two, a practically unlimited array of actions (Searle, 1965: 221). If (as is surely the case) one of the uses of language is to subordinate

another, what is interesting from a regulation perspective is the kind of subordination it is, and whether it ought to be regulated. That it is done as an illocutionary speech act is as uninformative as noting that to cause GBH to someone is an action – it is to generalise to the point it fails to distinguish the specific action from any other: to the point of dissolution. Langton argues that pornography is capable of engendering various *specific* harms and capable of performing the particular action of subordinating women:

> Speech can subordinate when it has a certain verdictive and exercitive force: when it unfairly ranks members of a social group as inferior, when it legitimates discriminatory behaviour towards them, and when it unjustly deprives them of important powers. (2009: 62)

Perhaps pornography does do these things as a separate question, and can have these forces, but as I argue her contention is not with whether or not the purported subordination done by pornography is an illocutionary act, but rather whether it is illocutionary in a specific way, or rather whether it 'does' something specifically regulable.

In addition to her argument regarding pornography, like MacKinnon, Langton extends it to hate speech. Drawing on Austin, she writes that "[i]n terms that J.L. Austin made famous, there appear to be both illocutionary and perlocutionary dimensions to hate speech" (2012: 75). This, again, is an unremarkable claim; hate speech, by virtue of being speech at all, has illocutionary and perlocutionary dimensions – it is used to perform certain actions and has certain effects, like any utterance does. In support of her observation that hate speech is an illocutionary act (and that this is legally interesting), she draws on the Nuremberg trials. She writes that "when Julius Streicher was tried […] His speech was an illocutionary act: he 'incited' his countrymen to persecute the Jews" (2012: 76). This is a clear example of the error in question. Because incitement to persecute is considered a particularly alarming and striking use of language, she places importance on the fact it was an illocutionary act generally. This places the cart before the horse: that *particular instance* of incitement was of interest for guilt at Nuremberg because of what it did, but not because the speech was an illocutionary act of incitement. The fact it 'acts' illocutionarily as a speech act reveals nothing about the reasons to regulate it, and the implication that it matters rests on the descriptive commitment to speech that is capable of being particularly illocutionary.

McGowan has developed the hate speech aspects of speech act theory considerably. She argues similarly for attention to understanding hate speech as action: "[r]ather than focus on the causal effects of racist hate speech (as the standard liberal defence does), I focus instead on the actions constituted by such speech" (2012: 122). She draws a distinction between her approach – focusing on the actions constituted by hate speech – with what she calls the standard liberal approach (which argues for regulation if speech has sufficiently harmful consequences):

> [A]ccording to the standard liberal approach, the harms *caused* by racist hate speech are outweighed by the harms that would be *caused* (to our system of free speech) were we to regulate it. Clearly, the focus is on the harms caused. (McGowan, 2012: 122)

She therefore distinguishes the argument that hate speech causes enough harm to regulate it with her argument which "is importantly different" (McGowan, 2012: 122), and she "instead focus[es] on what the speech in question does (illocutionarily or otherwise)" (McGowan, 2009: 122). This paper is not an endorsement or defence of what she refers to as the standard liberal approach, but rather a criticism of the proposed alternative. Rather than defend the liberal approach against which she contrasts her own argument, I instead here only go as far as to suggest that we should resist the claim that "some racist hate speech ought to be regulated because it constitutes an (otherwise illegal) act of racial discrimination" (McGowan, 2012: 122).

McGowan's theory is in part concerned with aspects of hate speech to which my critique here does not pertain. For instance, a significant aspect of her argument is that hate speech "can enact norms that prescribe harmful practices" (2019: 18). This is to say that some speech might cause a social environment to change by enacting norms – for instance, the use of a pejorative might enact a social norm that it is permissible to treat the target of such a pejorative in certain ways. It may be that some speech enacts norms that prescribe harmful practices and which warrants regulation, further discussion on which is considered elsewhere (Waldron, 2012; Barendt, 2019). As I argue though, this cannot be enhanced by the claim that this is done as an illocutionary speech act, or that it constitutes an act of racial discrimination. After all, McGowan acknowledges that "[c]onstituting harm […] is really just a very specific way of causing it; it is to cause it via adherence to norms enacted" (2019: 23). If enacting certain social norms causes harms that are worth regulating, this is one thing, as some speech might result in a form of discrimination or enact discriminatory norms which would deserve regulation, but, if I am correct in my clarifications as to the lessons of speech act theory, the fact it also *constitutes an act of* discrimination adds nothing to the strength of any arguments for or against regulation. McGowan's reliance on speech act theory to identify regulable illocutionary speech is contingent on her misconception that "Austin drew our philosophical attention to the fact that saying sometimes constitutes doing" (2005: 39). As illustrated in Section Two, it is widely held in speech act theory that saying *always* constitutes doing, which undermines the significance of the observation that hate speech, or any speech, is doing something as a speech act.

McGowan's argument for regulating hate speech is therefore at least in part that hate speech constitutes an act of racial discrimination, rather than speech which results in discrimination:

> [I]f some racist hate speech constitutes illegal acts of racial discrimination, then such utterances should not count as speech in the First Amendment sense and so the regulation of such (uncovered) utterances should raise no free speech concerns at all. (2012: 141)

This leads to her conclusion that "if my argument here is successful, some racist hate speech is not even within the scope of a free speech principle" (2012: 125). This flourish relies on the misapplication of speech act theory I identify in this paper. The tensions

between protecting free speech and regulating hate speech cannot be circumvented by this proposed method, despite McGowan's conclusion that:

> If it isn't even speech in the free speech sense, its regulation poses no such threat to our commitment to free speech [...] the way ought to be clear for the regulation of (some) racist hate speech in our public spaces. (2012: 145)

This threatens to sidestep the complex normative tensions between potentially harmful speech and free speech as a value and is conceptually fallacious. Even if some hate speech does things which are worth regulating against a commitment to free speech generally, this should be faced head-on rather than by re-identifying any such speech (arbitrarily) as an illocutionary act or a performative utterance in the ways explored in this paper. The danger of this misconception is that speech which may be especially vulnerable to regulation is dismissed as simply performative/illocutionary, and complex but necessary debate which might otherwise lead to the protection of the given speech is avoided. Crucially, as a method for identifying regulable speech, it confirms any prior bias we may have individually with respect to regulable speech in a way that appears to have categorical rigour behind it: speech which one finds particularly objectionable appears, on the surface, more conduct-like.

As one final contemporary example of the error in question, Maitra writes in a similar line to McGowan and Langton that "[t]o say that speech *constitutes* subordination is a different (and more controversial) claim than to say that it merely *causes* subordination" (2012: 96). She draws on Langton's work in making this argument: "Langton appeals to speech act theory, and in particular, to Austin's notion of an illocutionary act, to develop this constitutive claim" (2012: 98). Like Langton and McGowan, Maitra distinguishes her method of identifying regulable speech from a harm-based analysis which focuses on the effects of speech, writing that:

> [S]peech can *constitute* subordination [...] subordination is not merely a downstream causal consequence of the speech in question [...] under certain circumstances, the very act of producing speech of that kind just *is* a subordinating (speech) act. (2012: 98)

She offers what she takes to be a paradigm instance of an illocutionary act in a lawmaker during South African apartheid, arguing that speech used to enact a law which deprives people of rights "just *is* an act of subordination" (Maitra, 2012: 99). Given this example, she claims that there "is no conceptual difficulty with the claim that speech can constitute subordination [...] by constituting norms that help to construct reality for the subordinated group" (2012: 99). This observation is to note that some speech can construct certain legal norms which are subsequently relied on to marginalise a group harmed by these new norms. In the instance of relevant apartheid lawmakers, the construction of these norms through legal enactment is clear. She extends this observation to the potential regulation of some ordinary (non-law-making) hate speech, which she argues can have sufficient authority – differently but analogously (2012: 109). Therefore, in her analysis, ordinary hate speech can be an illocutionary act of subordination in the same way

laws that create discriminatory policy can be. However – crucially for my analysis – what is potentially interesting about this claim is whether this analogy holds and, if so, what it means for the regulation of any such speech, rather than its role as an illocutionary act which constitutes a regulable speech act.

My response to Maitra's line of argument does not differ in substance from my response to the other theorists I respond to in this paper and can be set in short form. She is concerned with a particular illocution speech has, which, normatively, may justify the regulation of that speech (contingent on her empirical claim that social linguistic norms can have similar authority to legal norms). Some forms of subordination that are performed via speech might warrant regulation – such as the enactment of some social norms, as she argues (2012: 96) – but this cannot be attributed to the fact that it "just *is* an act of subordination" (2012: 99).

Though the theorists examined in this Section make their arguments in different ways, they all share the fundamental premise – initially found in Greenawalt's and MacKinnon's work – of advocating for the regulation of pornography and hate speech on the basis that they are performing an act illocutionarily or performatively. It is not that these theorists are wrong that pornography or hate speech is an illocutionary speech act (or a performative utterance in MacKinnon's case), it is that it is an unremarkable and trivial claim that it is so. The illocutions involved in some speech may result in some social cost that the speaker ought to legally answer for – as a question of legal, moral and political value – or they may not, but the argument either way is not attributable to the fact they are illocutions or performatives. If the fact speech 'acts' erodes the line between hate speech and action, then it erodes the line between *all* speech and action, but not some select speech which also just so happens to be speech that one wants to otherwise regulate. This more general dissolution of the distinction between speech and conduct may be an argument some are willing to embrace as a separate but related issue (Harel 2011), but it is not possible to *isolate* any particular kind of speech within such a dissolution. If it is significant that speech constitutes illocutionary acts, it might be argued to be a universal erosion of the line between all speech and action, or no erosion at all. I do not engage with this alternative question here, but what can be said is that there is not *some* speech which can be likened to performing illocutionary speech acts (or performatives) in a special way. If a specific kind of speech action warrants regulation, then this argument should be made on the basis that it is something that deserves to be regulated, and these applications of speech act categorisation cannot be of aid in these specific value-led questions.

DOI: 10.14762/jll.2022.078

5. Conclusion

This paper has examined theories which invoke speech act theory to claim that speech might be better understood as regulable by attention to its illocutionary or performative nature. Greenawalt's early theory was in part constructed on the idea that some speech works as performative and enacts changes in the world by doing things. He applied his theory to a wide array of different uses of speech but the point was, distinctly but similarly, taken up in the context of hate speech and pornography – initially by MacKinnon and more recently by McGowan, Langton and Maitra. In this paper, I brought these accounts together to show their shared reliance on speech act theory in establishing the legal significance that a given utterance acts as either a performative or illocutionary utterance. These utterances, they all argue in their specific ways, do not merely say something, nor do they merely cause harmful effects which warrant regulation, but rather they constitute a form of speech action themselves. The consequences of their proposed way of examining speech regulation are significant. *If* it is right that (some) legally problematic speech is understood to be less to do with saying things and more to do with the performance of illocutionary speech acts, such speech can be, at a descriptive and objective level, understood to have less relevance for protection. The counterargument of this paper though has been that this regulatory attention to what speech performs or constitutes acts is unable to be supported by speech act theory as is claimed. This has been a mistaken assumption for some time, manifesting in different work, as this paper has sought to demonstrate.

It is inevitable that any legal system committed to free speech as an ideal will wrestle with competing values when it comes to what (and what not) to regulate. It is understandable that attempts to justify hate speech and pornography regulation particularly have taken this linguistic turn in pursuit of philosophical categorical assistance, but if the rectification I offer here is correct, then we need to face this challenge head-on in all its normative complexity without recourse to an erroneous taxonomy of speech which performs illocutionary or performative speech acts.

References

Adler, Amy (2001). Inverting the First Amendment. *University of Pennsylvania Law Review*, *149*(4), 973–1002.

Austin, John L. (1962). *How to Do Things With Words*. Oxford: University Press.

Austin, John L. (1979). Performance Utterances. In Urmson & Warnock (Eds.), *Philosophical Papers* (3rd ed.) (pp. 232–253). Clarendon Press.

Barendt, Eric (2019). What is the Harm of Hate Speech. *Ethical Theory and Moral Practice*, 22(2), 539–553. DOI: 10.1007/s10677-019-10002-0.

Butler, Judith (1997). *Excitable Speech: A Politics of the Performative*. New York/London: Routledge.

Cerf, Walter (1969). Critical Review of How to Do things with Words. In Fann (Ed.), *Symposium on J. L. Austin* (pp. 351–379). London: Routledge & Kegan Paul.

Collavin, Elena (2011). Speech Acts. In Bublitz & Norrick (Eds.), *Foundations of Pragmatics* (pp. 373–396). Berlin/New York: De Gruyter Mouton. DOI: 10.1515/9783110214260.373.

Gelber, Katharine (2012). 'Speaking Back': The Likely Fate of Hate Speech Policy in the United States and Australia. In I. Maitra & M. K. McGowan (Eds.), *Speech and Harm: Controversies Over Free Speech* (pp. 50–71). DOI: 10.1093/acprof:oso/9780199236282.003.0003.

Graham, Keith (1977). *J.L. Austin*. Harvard: University Press.

Greenawalt, Kent (2015). *Interpreting the Constitution*. Oxford: University Press.

Greenawalt, Kent (1989). *Speech, Crime and the Uses of Language*. Oxford: University Press.

Harel, Alon (2011). Is Pornography a Speech Act and Does it Matter? *Jerusalem Review of Legal Studies*, 3(1), 5–14. DOI: 10.1093/jrls/3.1.5.

Holdcroft, David (1978). *Words and Deeds: Problems in the Theory of Speech Acts*. Oxford: University Press.

Langton, Rae (2009). *Sexual Solopsism: Philosophical Essays on Pornography and Objectification*. Oxford: University Press. DOI: 10.1093/acprof:oso/9780199247066.001.0001.

Langton, Rae (2012). Beyond Belief. In Maitra & McGowan (Eds.), *Speech and Harm: Controversies Over Free Speech* (pp. 72–93). Oxford: University Press. DOI: 10.1093/acprof:oso/9780199236282.003.0004.

MacKinnon, Catharine (1993). *Only Words*. Harvard: University Press.

MacKinnon, Catharine (2012). Foreword. In Maitra & McGowan (Eds.), *Speech and Harm: Controversies Over Free Speech* (pp. vi–xviii). Oxford: University Press.

Maitra, Ishani (2012). Subordinating Speech. In Maitra & McGowan (Eds.), *Speech and Harm: Controversies Over Free Speech* (pp. 98–120). DOI: 10.1093/acprof:oso/9780199236282.003.0005.

McGowan, Mary K. (2005). On Pornography: MacKinnon, Speech Acts, and 'False' Construction. *Hypatia*, 20(3), 22–49. DOI: 10.1111/j.1527-2001.2005.tb00485.x.

McGowan, Mary K. (2012). On 'Whites Only' Signs and Racist Hate Speech: Verbal Acts of Racial Discrimination. In Maitra & McGowan (Eds.), *Speech and Harm: Controversies Over Free Speech* (pp. 121–147). DOI: 10.1093/acprof:oso/9780199236282.003.0006.

McGowan, Mary K. (2019). *Just Words: On Speech and Hidden Harm*. Oxford: University Press. DOI: 10.1093/oso/ 9780198829706.001.0001.

Schauer, Frederick (2015). On the Distinction Between Speech and Action. *Emory Law Journal*, 65(2), 427–454.

Searle, John R. (1965). What is a Speech Act. In Black (Ed.), *Philosophy in America* (pp. 221–239). Cornell: University Press.

Searle, John R. (1969). *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: University Press.

Tourkochoriti, Ioanna (2020). How Far Should the State Go to Counter Prejudice? A Positive State Obligation to Counter Dehumanisation. *Erasmus Law Review*, 3, 34–48. DOI: 10.5553/ELR.000162.

Waldron, Jeremy (2012). *The Harm in Hate Speech*. Harvard: University Press.

Weston, Daniel (2020). *Free Speech, Speech Act Theory and 'Speech Brigaded with Action': When Does Protected Speech Become Regulable Action?* [University of Leicester]. URL: leicester.figshare.com/articles/thesis/Free_Speech_Speech_Act_Theory_and_Speech_Brigaded_with_Action_When_Does_Protected_Speech_Become_Regulable_Action_/13289297 (accessed 3 October 2022).

Weston, Daniel (2022). A Model for Free Speech. *International Journal for the Semiotics of Law*, (Online First/Forthcoming). DOI: 10.1007/s11196-022-09918-1.

Wirenius, John F. (2004). *First Amendment, First Principles: Verbal Acts and Freedom of Speech*. New York: Holmes & Meier.

Yong, Caleb (2011). Does Freedom of Speech Include Hate Speech? *Res Publica*, 17(4), 385–403. DOI : 10.1007/s11158-011-9158-y.