Canadian Journal of Law & Jurisprudence XXIX No.2 August 2016, 457-494

© *Canadian Journal of Law & Jurisprudence* 2016 doi: 10.1017/cjlj.2016.20

Defining 'Speech': Subtraction, Addition, and Division

Robert Mark Simpson

1. Introduction: When is Speech, 'Speech'?

One of the basic tenets of liberalism is the idea that harmless conduct generally should not be subject to legal restriction. Liberal free speech principles go beyond this basic tenet. They demand that we accord a special status to 'speech', such that when it produces harm, it should be less liable to legal restriction than (similarly harmful) non-'speech'. Obviously, though, the million-dollar question is 'which communicative acts are appropriately classified as 'speech'?' And the answer plainly can't be 'whatever we call 'speech' in ordinary discourse', since many acts that involve 'speech' in the everyday sense of the word—like fraud, perjury, and extortion—fail to qualify as 'speech' under any plausible conception of free speech. Within free speech theory, then, 'speech' needs to be defined as term of art. But if free speech principles are to have any credibility, this term of art can't be an arbitrary contrivance; as Frederick Schauer says, it will need to be "defined by the purpose of a deep theory of freedom of speech".[1] In short, we require a definition of 'speech' that picks out all—and only—those acts that *really do* merit special protection, in light of whatever cogent arguments for free speech we can find. This paper is a critical inquiry into the methodologies that are employed in attempts to formulate such a definition of 'speech'. (Note that for the rest of this paper I'll be using the term 'SPEECH'—thus stylized in small capitals—to denote the putative class of behaviors that merit special protection against legal restriction under liberal free speech principles.)

Right from lesson one, students of free speech can see that SPEECH needs to be defined as a term of art. They see it in *On Liberty*, in Mill's claim that a statement of the opinion that corn dealers are starvers of the poor "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."[2] Students are reminded again that SPEECH is a term of art in Justice Holmes's famous remark, in his opinion in *Schenck v. United States*, that even

Thanks to the editor and an anonymous referee from this journal for their comments on an earlier draft. This paper has been improved by feedback from audiences at the University of Chicago Law and Philosophy Workshop, the Northwestern University Practical Philosophy Workshop, the 2015 Workshop on Global Expressive Rights at Dartmouth College, and the Melbourne Legal Theory Workshop at the University of Melbourne. Thanks to Susan Brison, Raff Donelson, Kath Gelber, Brian Leiter, and Lael Weis for these opportunities to present, and thanks especially to Natalie Stoljar and Adrienne Stone for preparing response presentations to my paper at Dartmouth and Melbourne respectively. For comments, criticisms, and suggestions, I'm grateful to all of the aforementioned people, as well as Ryan Doerfler, Amanda Greene, Andy Koppelman, Genevieve Lakier, Mary Kate McGowan, Martha Nussbaum, David Strauss, and Heather Whitney.

1. Frederick Schauer, *Free Speech: A Philosophical Inquiry* (Cambridge: Cambridge University Press, 1982) at 91.
2. John Stuart Mill, *On Liberty* (London: Penguin Books, 1985 [1859]) at 119.

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press

"the most stringent protection of free speech would not protect a man falsely shouting 'fire' in a theater".[3] Both examples show that free speech principles don't treat all 'speech' as 'SPEECH', and that they are only meant to constrain restrictions on the latter. All this is fine as far as it goes. But these examples also subtly tie free speech discourse to a further theoretical commitment, one which is rarely made explicit, let alone critically examined. On Holmes's view, shouting 'fire' shouldn't qualify as SPEECH, because it creates an imminent danger that the government has a legitimate interest in preventing. On Mill's view, the statement that corn dealers are starvers of the poor—if it is addressed an agitated mob—shouldn't qualify as SPEECH, because under these conditions this utterance constitutes a positive instigation to violence, and therefore constitutes a type of 'conduct' rather than SPEECH. In both scenarios what's suggested is not only that SPEECH is a special term of art, but also that the way to map-out the boundaries of this term is via a method of *subtraction*. As a default supposition, we assume that all verbal communications count as SPEECH, but then we make reasoned exceptions to weed out the various types of communicative *conduct* which do not in fact warrant the special protections accorded to SPEECH.

Granted, calling this a 'method' is a little overblown, since this suggests the deliberate application of a codified technique for defining SPEECH, whereas what we see in practice—in jurisprudential scholarship, in political philosophy, and in the writings of judges and legislators—is generally more haphazard. Many concerns and desiderata influence attempts to define SPEECH, and this results in candidate definitions that are often *ad hoc*, sometimes nakedly biased, and always contestable. Nevertheless, there is an underlying approach that's frequently in effect in efforts to define SPEECH, one that's prefigured in the examples that I've noted, and which all too often passes without scrutiny. We should seek to better understand this 'Subtractive Approach' to defining SPEECH, identify the form and complexion that it imparts to free speech theory, and think more about alternative definitional approaches and their comparative merits.

My discussion is organized as follows. In §2 I lay out some of the key assumptions about free speech theory that I'll be working with throughout. In §3 I briefly summarize how SPEECH is defined in U.S. First Amendment jurisprudence. In §4 I discuss an important critique of the way that SPEECH is defined for the First Amendment, which focuses on the politicized dimensions of this process, and I connect this critique with issues of definitional methodology. In §5 I discuss an alternative—more systematic and criterial—Subtractive Approach to defining SPEECH, which can be viewed as an attempt to remedy certain shortcomings of the U.S. approach. In §6 I explain why this criterial method is of limited benefit, since its usage requires interpretative judgments that give ready expression to the presumptions and biases whose effect we need to try to curtail in defining SPEECH. In §7 I discuss the most promising version of the Subtractive Approach, and note that it also has significant limitations. I then identify two wrongheaded ideas about what disqualifies communication from the domain of SPEECH that seem to be in effect across different versions of the Subtractive Approach to

3. 249 US 47 (1919).

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press

defining Speech. From this point onwards I consider theoretical alternatives to the Subtractive Approach. In §8 I explain why simply 'adding' things into the category of Speech, as an alternative to 'subtracting' them out, doesn't enable us to define Speech any better, unless we revise our conception of Speech so as to reject the two wrongheaded ideas from §7. The first one, which I elaborate on in §9, is that Speech is somehow inert, or of its essential character unlike 'real conduct'. The second one, which I discuss in §10, is that 'Speech' denotes a coherent and unified—rather than ad hoc and heterogenous—class of communicative behaviors. In §10 I present a sketch of a free speech theory that subdivides the category of Speech into several independent classes of protected communicative conduct, then does away with the umbrella category altogether. In §11 I explain why it isn't a problem that this kind of Subdividing Approach to free speech theory goes against free speech orthodoxy, by refusing to define Speech in a way that errs in favor of overprotection. In §12 I indicate some concerns, beyond bare theoretical curiosity, that motivate the development of this alternative Subdividing Approach to free speech. In §13 I sum up, then I defend the idealizing—which is to say, somewhat infeasible—nature of this kind of free speech theoretical framework.

2. The Specialness of Speech, and the Free Speech Arguments

As I say above, liberal free speech principles demand that Speech be treated as *special*, i.e., accorded protection against legal restriction that goes beyond liberalism's axiomatic injunction against the restriction of acts that don't harm others.[4] Given the harm principle, we don't need a separate free speech principle to articulate our opposition to the punishment of religious heresies, say, or the prohibition of 'decadent' art. Such things should be free from legal restriction simply because they are, in the relevant sense, *harmless*.[5] In cases where support for free speech is just opposition to these types of censorship, what's being supported is really just the rudiments of liberalism. A distinct principle of free speech goes beyond this. It demands that, with regards to the state's restriction of Speech that isn't harmless—Speech that causes or constitutes (or runs the risk of doing) genuine harm to others—a stricter justificatory standard for this

4. On this characterization of Speech as *special*, see Frederick Schauer, "Must Speech be Special?" (1983) 78:5 Nw UL Rev 1284; Douglas N Husak, "What is so Special about (Free) Speech?" (1985) 4:1 Law & Phil 1. In what follows I say nothing about non-verbal communicative acts, like the wearing of a black armband, being counted as Speech. I assume the standard view, that some such communicative acts can be categorized as Speech because they're communicative in a way that's relevantly similar to paradigmatic forms of Speech; see Paul Berckmans, "The Semantics of Symbolic Speech" (1997) 16:2 Law & Phil 145. In any case, this point has little bearing on an assessment of the Subtractive Approach and how it compares to other definitional methods. No method for defining Speech should categorize non-verbal communication as Speech by default, and any method can widen the borders of Speech to include some instances of non-verbal communication.

5. The relevant notion of harmlessness here is one that's correlative to a particular technical sense of 'harm', on which a harm is something like a rights-violating setback to an individual's interests; see for instance Joel Feinberg, *Harm to Others: The Moral Limits of the Criminal Law*, vol 1 (Oxford: Oxford University Press, 1985).

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press

restriction must be met, than in a case where equivalent harms stem from regular (i.e., non-SPEECH) action. A 'free speech theory', as I'll be using that expression, is a theory that specifies what that higher standard of justification for restricting SPEECH should involve, in principle and in practice, and which spells out the particular reasons why a higher standard of justification for restricting SPEECH is warranted, as opposed to just being something arbitrary or conventional.

Free speech theory, thus characterized, is an enterprise that only makes sense against the backdrop of liberalism, or a political philosophy that shares liberalism's relevant elements, i.e., its prioritization of the individual's liberty-protecting schedule of rights, and its corresponding commitment to the harm principle as a constraint on legitimate government action.[6] There may be a story to tell about what free speech looks like in a Marxist, Anarcho-syndicalist, Theocratic, or Communitarian Christian society, but 'free speech theory' here (and customarily) isn't that. There may also be a story to tell about free speech as a part of some substantial conception of human flourishing. If a person's views about the good life are all linked to open-minded inquiry and discussion then, for her, 'free speech' won't merely be about the side-constraints limiting legitimate government action, it will also be an affirmative ideal that influences her relationships, her vocation, and other major elements of her life. But free speech theory here (and customarily) isn't about all that; it's about a formal constraint that government must respect in order for it to be justified in restricting people's conduct, whether through the law or more directly.

If it is a mistake to regard SPEECH as special—if there are no good reasons to specially protect an elected class of communications, in a way that goes beyond liberalism's elementary injunction against restricting harmless acts—then the role of free speech ideals and principles in policy-making is (or should be) minimal.[7] To reach this conclusion is not to abandon liberalism wholesale. After all, there are at least some mature liberal democracies which, while nominally subscribing to free speech ideals, don't accord SPEECH any categorical privileges in their policy-making calculus. Under the Canadian Charter of Rights and Freedoms, for instance, there isn't any rigid injunction against restricting 'expression'. Such restrictions are permitted, under the Charter, provided that they are (in the relevant sense) 'reasonable' and 'justified'.[8] One result of this is that censorship disputes in Canada don't hinge on what gets categorized as 'expression'. This is why L.

6. Here and below I stress that liberalism begins with the harm principle, and that free speech principles don't do any distinctive work unless they operate against the backdrop of the harm principle. But this isn't meant to suggest that in the U.S. or any other liberal democracy the harm principle is enacted with anything like full consistency; after all, most liberal states prohibit harmless acts, e.g., related to public nudity. (Thanks to Martha Nussbaum for pressing me on this.) My point is that where there isn't even any notional allegiance to the harm principle—e.g., where heterodox ideas are routinely suppressed, without even any pretense that this is done for the sake of preventing harm to others—it's inapt to speak of violations of free speech. Free speech principles, properly characterized, are not merely one ramification of the harm principle, they are an independent supplement to it.

7. See Lawrence Alexander & Paul Horton, "The Impossibility of a Free Speech Principle" (1984) 78:5 Nw UL Rev 1319.

8. See §§1-2 of *The Canadian Charter of Rights and Freedoms*, Part I of the *Constitution Act* 1982, being schedule B to the *Canada Act 1982* (UK), 1982, c11.

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press

W. Sumner, in his analysis of anti-hate speech law in Canada, is willing to count any act with *any* expressive intent at all, even terror-bombing, as 'expression'.[9] If the question that guides policy-making is whether α's restriction is favored by a consequentialist calculus that encompasses all relevant values, it makes little difference to policy whether or not α is counted as 'expression' (or SPEECH). Things operate differently when we treat SPEECH as special, and in so doing demand that a stricter standard than mere 'reasonableness' or 'routine harm-prevention' be met in putative justifications for legally restricting SPEECH. At the same time, this doesn't mean that where SPEECH is accorded special privileges, policy is entirely *decided* by what gets classified as SPEECH. Where SPEECH is accorded a special status, to classify α as SPEECH isn't to disallow restrictions on α *regardless* of α's harmfulness.[10] Or in other words, not all conduct that is 'covered' by free speech principles is ultimately 'protected' against restriction.[11] What's more, the free speech tradition also contains a principle—namely, a general prohibition on government restricting any sort of conduct as a way of suppressing disapproved ideas—that can apply more or less independently of judgments about free speech coverage.[12] The point here is that giving a specially privileged status to SPEECH, as in the U.S., does have a considerable effect on censorship policy. Consequently, if we want to not downgrade the role free speech principles play in shaping policy, debates around free speech coverage cannot be ignored.

The question remains, as to why SPEECH *should* be deemed special. And in the free speech tradition—in political theory, philosophy, jurisprudence, and in First Amendment discourse—we find an oversupply of potential answers. Moreover, in addition to many different attempts to answer the first-order question, i.e., 'why should we specially privilege SPEECH', the literature also contains plenty of second-order analyses that try to summarize and systematize responses to the first-order question.[13] Now, the present work is about how SPEECH is defined in free speech, as distinct from what justifies free speech. The two things are related, but it's hard to say anything of substance about the justificatory questions without this becoming one's focus. This is why the following paragraph contains mostly assertions, rather than arguments.

There are four types of free speech justifications that are credible enough to merit serious attention. There is also a crucial 'meta-justification' that is

---

9. LW Sumner, *The Hateful and the Obscene: Studies in the Limits of Free Expression* (Toronto: University of Toronto Press, 2004) at 10.

10. At any rate, this is how we should understand the specialness of SPEECH in a legal system that purports to impose categorical constraints on government regulation of communicative action. Certain recent U.S. Supreme Court decisions cast doubt on whether this is the operative understanding of SPEECH's specialness under current First Amendment doctrine. For example, in *Snyder v Phelps* 562 US 443 (2011), the court appears to rule that it would be unconstitutional to restrict members of the Westboro Baptist Church from picketing soldiers' funerals *irrespective* of the pain that that activity causes to others; see Frederick Schauer, "Harm(s) and the First Amendment" (2011) 2011:1. Sup Ct Rev 81 at 87-90.

11. For instance, we might classify libel as SPEECH while at the same time regarding the harms caused or threatened by libeling SPEECH as sufficient, at least in principle, to justify its legal restriction; see Schauer, *supra* note 1 at 89-92.

12. Thanks to David Strauss for pressing me on this point.

13. See Schauer, *supra* note 1; Joshua Cohen, "Freedom of Expression" (1993) 22:3 Philosophy & Public Affairs 207.

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press

connected with each of the four. The first kind of free speech argument is *Epistemic*. Roughly, it says certain communication types—relating to inquiry, research, and reportage—must be specially protected, as SPEECH, since failing to do so will impair social epistemic goods.[14] The second kind of argument is *Democratic*. It says that certain communication types—relating to political advocacy, protest, and dissent—must be specially protected, as SPEECH, because failing to do so will undermine the legitimacy and/or the effectiveness of democratic politics.[15] The third kind of argument is *Expressive*. It says certain kinds of communications—e.g., relating to literature and the arts—need to be specially protected, as SPEECH, since failing to do so jeopardizes goods related to human creativity, authentic cultural expression, and aesthetic pleasure.[16] The fourth kind of argument pertains to *Worldview*. It says certain communications—relating to proselytism, ethical advocacy, and religious texts—must be specially protected, as SPEECH, in view of the need for all people to be able to give authentic expression to their ethical or spiritual belief-systems.[17] Each argument adverts to a particular aspect of human society, and each trains our focus on some specific types of communicative acts that it nominates for special legal protections. The meta-justification linking them all is one that alleges the incompetence or mendacity of *Government*. This argument is built around Libertarian concerns about government tyranny, of the kind that tend to recommend robust constraints on *all* government intervention in people's affairs, regardless of whether these aim at the restriction of communicative activity as such. In free speech theorizing, however, these Libertarian worries combine with the other types of arguments sketched above. So, for instance, it is the combination of (i) the fragility of certain social epistemic goods, and (ii) the mendacity and/or incompetence of government, that explains why social epistemic goods will be particularly imperiled without special protection for communicative conduct relating to research and reportage.[18, 19]

---

14. See, most famously, Mill, *supra* note 2.

15. See Alexander Meiklejohn, *Free Speech and its Relation to Self-Government* (New York: Harper & Brothers, 1948); or in late 20th century legal philosophy, Ronald Dworkin, *A Matter of Principle* (Oxford: Oxford University Press, 1985).

16. See, for instance, C Edwin Baker, *Human Liberty and Freedom of Speech* (Oxford: Oxford University Press, 1989); JM Coetzee, *Giving Offense: Essays on Censorship* (Chicago: University of Chicago Press, 1996).

17. See, for instance, Joseph Raz, "Free Expression and Personal Identification" in *Ethics in the Public Domain* (Oxford: Oxford University Press, 1994) 131.

18. Arguments from Government figure notably in Schauer's early free speech writing, especially *supra* note 1, and free speech theory regularly returns to them, especially in the work of American First Amendment scholars who find more utopian perspectives on free speech—as a royal road to truth or democracy—a bit too rose-colored; as in, for instance, Vincent Blasi, "The Pathological Perspective and the First Amendment" (1985) 85:3 Colum L Rev 449.

19. An argument that I'm ignoring in this section—one that's notable, but ultimately idiosyncratic—is the argument from tolerance, on which the reason for specially protecting SPEECH is that the state can thereby signal a tolerant ethos and thus help nurture tolerant sensibilities in society at large; see Lee C Bollinger, *The Tolerant Society* (Oxford: Oxford University Press, 1986). Another kind of argument I'm ignoring is the kind that stresses people's agency or autonomy as the key thing that's honored or promoted in a regime of communicative liberty; for a recent example see Seana Valentine Shiffrin, "A Thinker-based Approach to Freedom of Speech" (2011) 27:2 Const Commentary 283. Arguments of this latter kind have been roundly criticized; see, e.g., Susan J Brison, "The Autonomy Defense of Free Speech" (1998) 108:2

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press

How does all this link up with our methods for defining Speech? As I say in §1, a sound free speech theory will supply us with a definition of Speech which identifies all and only those communicative activities that really do merit 'special' protection, in the sense I have previously indicated. The arguments briefly sketched above, which are of course mapped-out in great detail in the literature, can inform us as to which kinds of communicative acts require such protection, and why. Henceforth I'll refer to these as 'the Arguments'. A successful definition of Speech won't simply piggy-back on our everyday usage of the term 'speech', or arbitrarily designate a class of communications with no evident relation to the Arguments. Instead, it will identify and specify unifying characteristics of the communication types which are indicated or nominated, by the Arguments, as being in need of special protection. When a definition of Speech satisfies this condition, I'll say that it 'answers to the Arguments'. An assessment of the Subtractive Approach to defining Speech is, first and foremost, an investigation into whether this method can yield a definition of Speech that satisfactorily answers to the Arguments.[20]

3. Defining Speech in the First Amendment

Although our subject is free speech *per se*, rather than the institutionalized form of free speech we find in U.S. constitutional law, it still behooves us to pay attention to the First Amendment, not just because it's the most comprehensive jurisprudential system of free speech principles, but also because it's a legal framework which, to a greater degree than any other, purports to make good on the basic idea that Speech should be specially privileged. Whole books could be written on how Speech is defined within the First Amendment, so my discussion here will be cursory. What I'm trying to do is (i) motivate a particular critique of how Speech is defined for First Amendment purposes, and (ii) show how one of the assumptions of the Subtractive Approach is sometimes at work in the circumscription of Speech for First Amendment purposes.

One significant area of First Amendment doctrine, vis-à-vis the boundaries of Speech, relates to so-called 'fighting words'. In *Chaplinsky v. New Hampshire* (1942), the Supreme Court upheld a New Hampshire statute under which Walter Chaplinsky had been convicted for calling a police officer a 'damned Fascist' and 'damned racketeer'. Justice Murphy's opinion, affirming the statute's

Ethics 312. I'm ignoring them not for that reason, but just because I think they advert to values that are accounted for in the other free speech arguments noted above.

20. In thinking critically about our methods for defining Speech, in accordance with the approach that I've outlined here, the wider theoretical aim is not merely to say what the operative conception of Speech in free speech theory and practice *is*, as a matter of contingent sociological fact, but rather to say what conception of Speech ultimately should be at work in free speech theory. In her influential work in social theory Haslanger calls this mode of inquiry into a contested concept an 'ameliorative inquiry'; see, e.g., Sally Haslanger, "Gender and Race: (What) are They? (What) Do We Want Them to Be?" (2000) 34:1 Noûs 31. For a detailed defense of the view that ameliorative inquiries are a legitimate part of conceptual analysis in law and legal theory, see, e.g., Natalie Stoljar, "What Do We Want Law to Be? Philosophical Analysis and the Concept of Law" in Wil Waluchow & Stefan Sciaraffa, eds, *Philosophical Foundations of the Nature of Law* (Oxford: Oxford University Press, 2013) 230.

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press

constitutionality, identified fighting words as a type of communication that's outside the protected realm of SPEECH. "There are certain well defined and narrowly limited classes of speech" he said, whose restriction has "never been thought to raise any Constitutional problem."

> These include the lewd and obscene, the profane, the libelous, and the insulting or "fighting" words—those which, by their very utterance, inflict injury or tend to incite an immediate breach of the peace … such utterances are no essential part of any exposition of ideas, and are of such slight social value as a step to truth that any benefit … from them is clearly outweighed by the social interest in order and morality.[21]

The chief consideration cited here to explain why fighting words don't count as SPEECH is (i) that they make a negligible contribution to social epistemic goods. Three further features of fighting words are mentioned, which can also be read as a part of the court's reasons for subtracting them from the domain of SPEECH: (ii) that fighting words injure by their very utterance, (iii) that they incite immediate disorder, and (iv) that the courts have never counted them as SPEECH. Now, at least some of these sound like *pro tanto* good reasons to place fighting words outside the domain of SPEECH. But at the same time, this reasoning looks like an improvised response to the issues created by fighting words policy, rather than an application of a general principle for classifying SPEECH, which has been settled in advance and disinterestedly applied to the case. We aren't told whether any of criteria (i)-(iv) are sufficient to disqualify speech as SPEECH, or whether all four are required, nor are we told about their relative priority, nor whether there are other conditions that disqualify communications as SPEECH, but which happen not to be salient vis-à-vis fighting words. *Chaplinsky* doesn't address these questions because it isn't founded upon a deeper theory that answers them. This lack of a settled theory of the boundaries of SPEECH is evident again in subsequent decisions around fighting words. *Cohen v. California* (1971),[22] for instance, rules that utterances, in order to be subtracted from SPEECH as fighting words, need to express not just general hostile sentiments, but "personally abusive epithets"; thus extending protection to the message on Cohen's jacket ('fuck the draft'), for which he had previously been convicted under the California penal code. Are such doctrinal tweaks pure *ad-hoccery*? Not necessarily. In this policy area, as in other areas of constitutional jurisprudence, the Court is attempting to discern the proper purposes of the First Amendment, and to ensure that protections are only extended to communicative conduct in a way that's justified in view of those purposes.[23] What this generates, though, are ways of thinking about the boundaries of SPEECH that

21. 315 US 568 (1942).
22. 403 US 15 (1971).
23. Outside the framework of legal institutions, some purely political-theoretical inquiries into what should be counted as SPEECH work at this same intersection, between the underlying normative aims of free speech principles and the social issues raised by a particular type of communicative conduct. One recent example of this, which examines whether hate speech should qualify as SPEECH, is in Caleb Yong, "Does Freedom of Speech Include Hate Speech?" (2011) 17:4 Res Publica 385.

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press

are relevant to a specific policy issue, rather than generalizable criteria which demarcate the boundaries of SPEECH.

Fighting words doctrine largely belongs to the First Amendment's past.[24] Questions about the status of corporate electioneering communication, by contrast, belong to its present, and a similar set of issues arises in this arena. In *Citizens United v. Federal Election Committee* (2010),[25] the Court's ruling—contrary to sections of the 2002 Bipartisan Campaign Reform Act (BCRA),[26] and to earlier Court decisions, including *Austin v. Michigan Chamber of Commerce* (1990)[27]—was that 'independent' expenditures by corporations, advocating the election or defeat of a candidate, *do* count as SPEECH, and should be protected from legal restriction accordingly.[28] *Citizens United* thus overrules prior judgments, by the courts and the federal legislature, which had subtracted such corporate electioneering from the protected domain of SPEECH. In Justice Kennedy's opinion for the (5-4) majority, this reclassification of corporate electioneering as SPEECH was defended via criticism of the validity of the reasons for which it had previously been subtracted. Corporate electioneering may have a distorting impact upon our public debates surrounding electoral politics, and it's fairly probable that allowing corporations to engage in electioneering communication on similar terms to labor unions will enable corporate interests to exert greater influence than organized labor on election outcomes. But corporate electioneering is political speech despite these things, so Kennedy argues, and the BCRA 2002 and *Austin* were mistaken to regard the above considerations as reasons to subtract it from the domain of SPEECH. As with fighting words doctrine, the reasoning that guides this classificatory judgment is related to the specific policy questions at issue, and therefore we aren't being presented with any fully generalized theory of what the First Amendment covers. But, as in the previous case, it would be wrong to deride the reasoning as pure *ad hocery*, even if we're dubious about its substance.[29] The court is purporting to honor the First Amendment's aims, and—in a way that's compatible with inherited doctrinal understandings of how those aims then translate into specific constraints upon government action—ensure that the forms of corporate electioneering which are encompassed within the First Amendment's aims receive the protections that they're entitled to.

Other areas of First Amendment doctrine could be examined to shed more light on our picture of the process by which SPEECH is defined under U.S. constitutional

---

24. The Court has narrowed its definition of fighting words each time it has applied it, and hence Chaplinsky himself, as Lawrence wryly notes, is "the only defendant … ever found to have used fighting words"; see Frederick M Lawrence, "Violence-conducive Speech: Punishable Verbal Assault or Protected Political Speech?" in David Kretzmer & Francine Kershman Hazan, eds, *Freedom of Speech and Incitement Against Democracy* (London: Kluwer Law International, 2000) 11 at 28.

25. 8-205, 558 US 310 (2010).

26. *Bipartisan Campaign Reform Act of 2002*, Pub L No 107-155, 116 Stat 81.

27. 494 US 652 (1990).

28. By 'protected', here, I mean 'subject to strict scrutiny in the process of judicial review'.

29. For critiques of the substance of the majority reasoning in *Citizens United*, *supra* note 25, see, e.g., Lawrence Lessig, "A Reply to Professor Hasen" (2012) 126 Harv L Rev F 61; Steven J Heyman, "The Conservative-Libertarian Turn in First Amendment Jurisprudence" (2014) 117:1 W Va L Rev 231.

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press

law, but these two examples will suffice for present purposes. These examples indicate two things. First, they show the Court's increasing reluctance, over time, to allow for new subtractions, or even just to maintain previous subtractions, of particular sub-categories of communicative conduct from the privileged domain of SPEECH.[30] In the present U.S. lawmaking environment, the onus of justification falls on those who think that a particular type of communication ought to be subtracted from the privileged domain of SPEECH, and thus placed outside the normative arena governed by First Amendment doctrine. What's more, even with a category of communication like 'obscenity', i.e., one whose subtraction from the domain of SPEECH has been a longstanding element of First Amendment doctrine, judicial interpretations of where SPEECH ends and obscenity begins have sometimes bestowed First Amendment protections on communicative acts that would have otherwise been unprotected.[31] It isn't a matter of historical necessity that things should work like this. Instead of a general presumption that communicative acts qualify as SPEECH, unless or until an analysis of the aims of the First Amendment forces us to conclude otherwise, it's at least possible that an alternative classificatory system could have evolved, under which all types of acts are located outside the technically-specified domain of SPEECH, until an analysis of First Amendment aims forces us to include them. In any case, the onus for now runs the opposite way.[32]

---

30. As one popular First Amendment casebook observes, "the court has not upheld a restriction on speech because it might induce readers or listeners to engage in criminal activity since Dennis (1951) …. The Court has not upheld a restriction on speech because it might provoke a hostile audience response since Feiner (1951). It has never upheld a restriction on speech because the ideas expressed might have an improper influence on the judicial process"; Geoffrey R Stone, Louis M Seidman, Cass R Sunstein, Mark V Tushnet & Pamela S Karlan, *The First Amendment*, 4th ed (New York: Wolters Kluwer Law & Business, 2012) at 125.

31. Obscenity is a rather complex case, though. The standard of obscenity outlined in *Roth v United States*, 354 US 476 (1957) was altered sixteen years later in *Miller v California,* 413 US 15 (1973), such that, in order to be prohibited, a work only needed to be found to lack 'serious' value, rather than it having to be found to be 'utterly without socially redeeming value'. *Miller* thus extended the compass of 'obscenity', and thereby—against the overriding trend in First Amendment coverage through this era—subtracted some communicative acts from the protected domain of SPEECH. But in between *Roth* and *Miller*, *Stanley v Georgia*, 394 US 557 (1969) included a clarifying ruling to the effect that the First Amendment disallows any prohibition on the mere private (i.e., non-commercial) possession and consumption of obscene material. In that ruling, *Stanley* effectively narrowed the range of communications held outside the domain of First Amendment protection via the excepted category of obscenity.

32. The brevity of my discussion here makes it difficult to avoid some oversimplification. As Schauer has recently reemphasized, there are some kinds of communicative behaviors (like perjury) for which there's never been *any* onus on *anyone* to justify their exclusion from SPEECH—which are, rather, simply universally accepted as being outside the domain of SPEECH, despite the fact that they obviously involve 'speech' in the everyday sense of the word; see Frederick Schauer, "Out of Range: On Patently Uncovered Speech" (2015) 128 Harv L Rev Forum 346. (Thanks to Andy Koppelman for pressing me on this.) To be clear, then, my remarks above primarily apply to types of communication whose status as SPEECH or non-SPEECH has been a matter of live controversy. I should also note that when it comes to questions of coverage, the historical evolution of First Amendment doctrine has been considerably less tidy than what's suggested in the broad-strokes picture that I've been painting. Lakier convincingly argues that things like libel and commercial advertising have had a volatile and unsettled trajectory—vis-à-vis their status as SPEECH or non-SPEECH—more so than is usually acknowledged in contemporary First Amendment jurisprudence; see Genevieve Lakier, "The Invention of Low-value Speech" (2015) 128:8 Harv L Rev 2166.

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press

The second thing these examples indicate is that judgments about when First Amendment doctrine should apply are not typically determined by an appeal to some prior, independent, and settled characterization of Speech. Indeed, something like the reverse is closer to the mark: what gets characterized as Speech is typically determined by judgments about when First Amendment doctrine should apply. And these judgments, about when the First Amendment is in play, tend to be issue-specific, since they partly hinge on narrow, imminent, practical questions (e.g., 'what problems does this statute aim to address?'), and not just the larger normative and interpretative questions about the First Amendment's purposes. For each type of communicative act at issue, courts have to consider the social impact of these acts, and the reasons why they're performed, and determine whether there is a good case for prohibition or regulation, but also whether this is overridden by some loftier reason to protect the communicative activity in question, related to the purposes the First Amendment is meant to serve. Justice John Paul Stevens spoke to this complexity in discussing the constitutional status of anti-hate speech laws. "Whether a particular act or message is more appropriately deemed 'speech' or 'conduct'" he says, "turns on context as well as content … a cross-burning as part of a public rally in a stadium may fairly be described as protected speech, [but] burning the same cross on the front lawn of an unfriendly neighbor has an entirely different character".[33] My point is that rulings which serve to demarcate the boundaries of Speech for First Amendment purposes are enacted amid the hurly-burly of these pragmatic-interpretative challenges. What we patently *don't* see, is a systematic method for distinguishing Speech from non-Speech, which can be used procedurally to decide when the First Amendment apparatus should be used to assess the regulation of a particular type of communication.

4. The Politics of Defining Speech

How Speech is defined for First Amendment purposes is a politically controversial matter to say the least. Some critics argue that Speech gets defined in a way that unjustifiably privileges the communicative aims of entrenched ideological interests. Part of this complaint is that both sides of politics use free speech rhetoric to shield their favored communications from legal challenge, while also contesting each other's rhetorical maneuvers to the same effect.[34] But the deeper worry is that Speech is—can only *ever* be—a makeshift category, and that what gets put into it, and thus specially protected, isn't a matter of principle, but just a result of some faction having the savvy to commandeer the First Amendment to further its own aims. As Stanley Fish says:

> Decisions about what is and what is not protected in the realm of expression will rest not on principle or firm doctrine but on the ability of some persons to interpret

33. John Paul Stevens, "The Freedom of Speech" (1993) 102:6 Yale LJ 1293 at 1310-11.
34. See Nat Hentoff, *Free Speech for Me But Not for Thee: How the American Left and Right Relentlessly Censor Each Other* (New York: Harper Collins, 1992).

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press

> … doctrine in ways that lead to the protection of speech they want heard and the regulation of speech they want silenced …. When the First Amendment is successfully invoked, the result is not a victory for free speech in the face of … politics, but a *political victory* won by the party that has managed to wrap its agenda in the mantle of free speech.[35]

Schauer has written on the same theme, though in a less iconoclastic mood than Fish.[36] It isn't formal legal precepts or abstruse free speech theory, he says, that determine why some issues are regarded as hinging on First Amendment doctrine, while others are handled in a way that sets First Amendment doctrine aside. The question of which acts are such that their restriction is a First Amendment issue is influenced, and often decided, by cultural and economic factors, which are of course highly politicized, and can easily become battlegrounds in culture wars.

This worry isn't merely academic. When legislation aimed at remedying the sexist effects of pornography, and campus speech codes aimed at protecting students from racist harassment, are deemed unconstitutional on First Amendment grounds, the characterization of these policies as First Amendment issues is not the upshot of a disinterested classificatory process, but rather a consequence of an unacknowledged ideological creed: that the liberties of sexists and racists matter more than the equality of women and minorities. At any rate, so feminist critics[37] and critical race theorists[38] have argued. And even if one isn't fully persuaded that the issues are ideologically rigged in the way these critics say, one can agree with the critics this far: that free speech principles don't merit any support *unless they really are principles*, as opposed to the hollow rhetorical devices that Fish bemoans. Theorists of a Foucauldian bent, for whom liberal ideals ultimately conceal structural oppression, will see little promise in trying to do anything with 'free speech' beyond unmasking the hidden currents of power it legitimizes.[39] But there is another way forward, if we're worried about a potentially malign politics of free speech, and how definitions of SPEECH are involved in this, which is to try to *recover* liberalism's egalitarian and emancipatory aims, and refashion liberal principles in their service.[40] Indeed, Fish points the way. In

---

35. Stanley Fish, *There's No Such Thing as Free Speech, and it's a Good Thing Too* (Oxford: Oxford University Press, 1994) at 110.

36. See Frederick Schauer, "The Ontology of Censorship" in Robert C Post, ed, *Censorship and Silencing: Practices of Cultural Regulation* (Los Angeles: Getty Research Institute for the History of Art and the Humanities, 1998) 147; "The Boundaries of the First Amendment: a Preliminary Exploration of Constitutional Salience" (2004) 117:6 Harv L Rev 1765.

37. See Andrea Dworkin & Catharine A MacKinnon, *Pornography and Civil Rights: A New Day for Women's Equality* (Minneapolis: Organizing against Pornography, 1988).

38. Richard Delgado, "Words that Wound: a Tort Action for Racial Insults, Epithets, and Name-calling" (1982) 17:1 Harv CR-CLL Rev 133.

39. Butler is one prominent contemporary figure who leans this way; see Judith Butler, "Ruled Out: Vocabularies of the Censor" in Robert C Post, ed, *Censorship and Silencing: Practices of Cultural Regulation* (Los Angeles: Getty Research Institute for the History of Art and the Humanities, 1998) 247.

40. There is some reason to think this perspective would have the support of some influential contemporary political philosophers who write on structural racial oppression. For instance, Mills's work on radical black liberalism explicitly calls for the recovery of liberal ideals in the name of a radical egalitarian politics; see Charles W Mills, "Occupy Liberalism! Or, Ten

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press

the politics of free speech, he says "it is not that there are no choices to make or means of making them", rather, it's that these choices "are inextricable from the din and confusion of partisan struggle".[41] Our aim shouldn't be to define Speech in a way that's neutral and apolitical, but to define it in a way that answers to the ideals that free speech politics in its classical, Millian guise—siding *with* the marginalized *against* the powerful—was supposed to serve.[42]

So what exactly would this lead us towards, methodologically? Here's a *prima facie* promising option: we need a method for defining Speech in which the crucial distinctions are made outside of the tumultuous intersection of policy-making and constitutional interpretation, so that they can be more firmly rooted in a clear and faithful conception of free speech's political aims, than is possible in that chaotic arena. Speech has to be defined by the purpose of a theory of freedom of speech, but this means defining it using abstract, generalizable criteria, which can differentiate the communicative acts that *really do* require special protection against legal restriction from the ones that don't. In §§5-6 I'll examine Kent Greenawalt's attempt, in *Speech Crime and the Uses of Language*,[43] to develop this sort of systematized, subtractive method for categorizing Speech, and I'll also look at Ishani Maitra and Mary Kate McGowan's structurally similar proposal, from two more recent papers,[44] which they use to argue that (some) hate speech and pornography shouldn't be counted as Speech. An understanding of how Speech is defined within the First Amendment (§3), and of the political issues around this (§4), helps us see what's appealing about these accounts, but also why they ultimately can't take us very far.

5. A Systematic Subtractive Method for Defining Speech

Greenawalt starts with an initial working supposition, that all 'oral and written communication' and 'nonverbal artistic endeavors' qualify as Speech. But soon thereafter he notes the obvious problem, that "some ordinary communications are reached barely, if at all, by the justifications for free speech".[45] The question is how to identify these acts, without relying on erratic intuitions about 'the point' of this type of utterance or that type of writing. Greenawalt's proposal is to subtract, from the domain of Speech, 'situation-altering' acts, i.e., acts which "change the social world in which we live", by "shifting rights or obligations or both".[46] Consider the example of a statement that constitutes a promise. When

Reasons Why Liberalism Cannot be Retrieved for Radicalism (And Why They're All Wrong)" (2012) 15:2 Radical Philosophy Review 305.

41. Fish, *supra* note 35 at 115.

42. On this point see Andrew Koppelman, "Veil of Ignorance: Tunnel Constructivism in Free Speech Theory" (2013) 107:2 Nw UL Rev 647.

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

44. Ishani Maitra & Mary Kate McGowan, "The Limits of Free Speech: Pornography and the Question of Coverage" (2007) 13:1 Leg Theory 41; "On Racist Hate Speech and the Scope of a Free Speech Principle" (2010) 23:2 Can JL & Jur 343.

45. Greenawalt, *supra* note 43 at 41.

46. *Ibid* at 58-59.

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press

person A says 'I promise to φ', she doesn't report a preexisting fact about her being obliged to φ; rather, A's utterance generates an obligation for her to φ, thus altering the social world that she and others inhabit. This is the sense in which "I promise to φ" qualifies as a situation-altering utterance under Greenawalt's account. More specifically, Greenawalt wants to subtract from SPEECH utterances that are situation-altering both 'substantially' (i.e., they effect substantial rather than negligible changes in duties or rights) and 'dominantly' (i.e., an alteration of the situation is the act's primary aim or effect).

The key to making sense of this terminology is to think of Greenawalt as attempting to reinstate an intuitive distinction between (i) words being used merely to *say* things, and (ii) words being used to *do* things. I say 'reinstate', because the viability of any such distinction has been dubious ever since the emergence of speech-act-theoretic work in linguistic pragmatics, popularized by midcentury philosophers like Austin, Strawson, and Searle.[47] The enduring insight of these early speech act analyses of language is that linguistic acts (whether written or spoken) cannot be divided into two discrete, non-overlapping classes of sayings and doings. Rather, all communicative, linguistic acts are both sayings *and* doings.[48] Or to use the jargon: all speech has a 'locutionary' aspect, i.e., it conveys certain content, and an 'illocutionary' aspect, i.e., its performance constitutes an act of some kind, whether it's a representative ('α is the case'), an expressive ('I'm sorry about α'), a declarative ('I pronounce thee α'), a directive ('you, φ, now!'), or a commissive ('I promise to φ').[49] Greenawalt recognizes that speech-act analyses undermine a SPEECH v. conduct binary. Nonetheless, he thinks there *is* a normatively significant distinction in the offing, that tracks this (once-) commonsensical binary, and which free speech principles are meant to latch onto. There is, he thinks, a primary purpose that can be ascribed to communicative acts: some of them aim primarily to convey ideas, opinions, information, claims, and putative facts, while others aim primarily to effect changes in our social relations. The former is what free speech principles are meant to protect, he maintains, and the criterion of being 'dominantly and substantially situation-altering' is meant to identify the latter, so as to exclude them from the protected domain of SPEECH. What this gives us, in theory, is a principled explanation of

47. JL Austin, *How to Do Things with Words* (Cambridge: Harvard University Press, 1962); PF Strawson, "Intention and Convention in Speech Acts" (1964) 73:4 The Philosophical Review 439; John R Searle, *Speech Acts: An Essay in the Philosophy of Language* (Cambridge: Cambridge University Press, 1969).

48. This will come as a surprise to those who read the initial, conjectural 'performative' versus 'constative' distinction in the early chapters of Austin, *supra* note 47, as the author's actual position, ignoring the tripartite taxonomy he develops in the later chapters, after having spelled out the various inadequacies of the performative versus constative distinction.

49. Different theorists propose different taxonomical sub-classes of illocutionary action; these ones are from Searle; see John R Searle, "A Classification of Illocutionary Acts" (1976) 5:1 Language in Society 1. The other dimension of speech acts, which I've set aside here, is the 'perlocutionary'. Words—as well as conveying content, and constituting acts by their very utterance—also produce effects. "Careful; it's raining outside" conveys the locutionary content that it's raining, and has the illocutionary force of a warning, but it also has the perlocutionary effect of, say, persuading someone to grab an umbrella. The perlocutionary act of persuasion isn't constituted by the utterance of the words (as an illocutionary act is), but rather effected by the utterance.

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press

why communicative act types such as agreements, promises, threats, and orders aren't all covered by free speech. And the account is then supposed to enable us to extrapolate beyond these cases, which are already widely reflected in law, so that judgments about whether, e.g., fighting words or corporate electioneering qualify as SPEECH can be made by us asking whether acts of these kinds are dominantly and substantially situation-altering.

Maitra's and McGowan's account is inspired by Greenawalt's, and the core idea is much the same: some communicative acts are more like doings than sayings, and we can differentiate these from SPEECH-proper by identifying their defining criteria *qua* doings. Greenawalt's talk of 'situation-altering' utterances is replaced with talk of 'obligation-enacting' utterances in their account, primarily because the latter term singles-out the type of change to social relations that matters most, namely, ones in which people's obligations are altered. The other, more important, difference between the two accounts is that Maitra and McGowan don't regard dominance, in Greenawalt's sense, as having a proper bearing on whether communicative acts should qualify as SPEECH. As long as a communicative act is *significantly* obligation-enacting, they say, it should be placed outside the domain of SPEECH, even if obligation-enactment isn't its primary aim or effect. What's the upshot of this change? Suppose person A makes a racist remark, and that this enacts social obligations in the context of utterance (e.g., by activating latent racial hierarchies), but suppose also that this obligation-enactment isn't the primary aim or effect of A's remark. By getting rid of the 'dominance' criterion, Maitra and McGowan want to allow, pace Greenawalt, that an utterance like this should be located outside the boundaries of SPEECH.[50]

Here is how this type of method for defining SPEECH would operate in theory. Suppose we're considering a proposal to regulate some form of apparently harmful communicative action, e.g., Homophobic Hellfire Preaching, or HHP.[51] Under the approach we see in First Amendment theory, the question of whether HHP qualifies as SPEECH is answered by asking whether, and to what degree, this type of communication is implicated in the goods and ideals in light of which a higher standard of justification in general has to be satisfied in order for SPEECH to be justifiably restricted. Under Greenawalt's approach to defining SPEECH, by contrast, we don't cross-check HHP with our justifications for free speech directly, like this; instead, we classify HHP as either SPEECH or non-SPEECH depending on whether it satisfies our abstract sorting criterion, e.g., for Greenawalt, being 'substantially and dominantly situation altering'. The free speech ARGUMENTS still play a part in this, insofar as the abstract sorting criterion is supposed to be crafted to track the crucial distinction, between communicative acts that merit special protection, *qua* free speech, and those that don't. But because these classificatory judgments are rendered using an abstract criterion, they are more mechanical, less beholden to prior prejudices about what should and shouldn't be counted as

50. Maitra & McGowan, *supra* note 44 (2007) at 54.

51. What I have in mind here are sermons delivered to church congregations, which espouse a deeply heteronormative sexual morality, and a doctrine of salvation that asserts a causal connection between sexual impurity and damnation.

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press

Speech, and less susceptible to the kind of *ad hoc* justificatory shoehorning that we're drawn into, when asking a question like 'does HHP partake of the ideal of democratic participation?', or any parallel question about whether some malign form of communication connects-up in any manner with any of the justificatory ideals that free speech Arguments invoke.[52] The merit of Greenawalt's method—the hope behind it, at least—is that by making these judgments abstract and criterial, it's easier for us to apply the whole free speech apparatus in only those cases that really do trigger the relevant normative concerns.

6. Substantiality, Dominance, and Significance

One odd thing about Greenawalt's approach to reinstating a speech-conduct distinction is its implied notion that Speech—i.e., communicative acts that aren't substantially situation-altering—doesn't succeed in 'changing the social world'. Granted, claims of opinion or putative fact don't alter agents' responsibilities and reasons in precisely the same way that promises and threats do. Nevertheless, for anyone engaged in scholarly research and writing, it seems strange—verging on a kind of performative contradiction—to say that *bona fide* Speech differs from mere verbal conduct by dint of its being inert with respect to the social world. One needn't be a radical linguistic constructivist in order to think that many types of Speech, even ones that don't fulfil the 'substantially situation-altering act' criterion, nonetheless *do* have a significant influence in shaping the social world. This is an issue I'll return to in §9. For our purposes here, the deeper problem with Greenawalt's approach is its inability to determine what counts as a 'substantially situation-altering act' when dealing with hard or contested cases. The provisional conclusion we arrived at in §4 was that Speech needs to be defined with reference to free speech's underlying purposes, but that this is best achieved via some kind of abstract, generalizable method, which can be used to systematically distinguish the communicative acts that really do merit protection against legal restriction from the ones that don't. This is what Greenawalt tries to provide. But when we look at the criteria he proposes for making the requisite systematic distinctions, we see that they require too much interpretative discretion in their application for them to informatively *guide* judgments about which types of communicative conduct are to be sorted into which class. Return to the example of Homophobic Hellfire Preaching. Is HHP dominantly and substantially situation-altering? Well, suppose one is a gay 15 year-old, dragooned into attending church every Sunday, where one is presented with heated sermons which tell one—purportedly with authority—that one's sexual desires render one worthy of damnation, and will result in precisely that if acted upon.[53] Is such communication situation-altering? Does it change

52. As in the way Dworkin strains to situate pornography within the protective domain of 'free speech'; see Ronald Dworkin, "Is There a Right to Pornography?" (1981) 1:2 Oxford J Leg Stud 177.

53. I'm not claiming that this is the 'correct', or most biblically accurate, or most charitable view of Christian sexual morality and soteriology (or that Christianity is worse than other major

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press

the social world one lives in, by shifting rights or obligations or both? And if we were to say that HHP *is* situation-altering, the next question is: is it *dominantly* so (is this the primary aim or effect of the speech) and *substantially* so (are the shifts in obligations substantial)? These are awkward questions. On one hand, the fact that HHP is an expression of religious convictions and ethical opinions suggests that—according to conventional ideas about free speech coverage—it *uncontroversially* qualifies as SPEECH, and merits protection against legal restriction accordingly. On the other hand, the fact that HHP—for some adolescents who are immersed in evangelical communities—will engender self-hatred, legitimize practices of violent bullying, and support noxious social hierarchies, gives us compelling reasons to judge that HHP is dominantly and substantially situation-altering, and thus to subtract it from the protected domain of SPEECH. I won't press the argument further either way. My point is about Greenawalt's method's inability to offer guidance in thinking about hard cases like these. It can't provide guidance in judgments about what counts as SPEECH, because all the key terms that it employs in its sorting criteria—'dominant', 'substantial', 'obligations', and 'rights'—admit of multiple interpretations, so that we're left with little choice but to look beyond these criteria—either to independent judgments about what SPEECH should encompass, or else to independent ethical judgements about what HHP does and how the law should respond to it—in order to resolve our classificatory quandary. The hope of the Greenawalt method was that these judgments could be regimented, and relocated to a deliberative space within which prejudices, pressures, and partisanship would be less likely to cloud our vision about what a free speech framework should be and should aim for. But it appears doubtful that they can.

Greenawalt defends his method by claiming that it vindicates prior judgments about what should and shouldn't count as SPEECH. Most jurisdictions that endorse free speech are already committed to the idea that verbal threats, perjuries, and contracts don't qualify as SPEECH for the purposes of free speech. Greenawalt hopes to *explain* why these judgments run the way they do, namely, because all these types of communication are dominantly and substantially situation-altering. Maitra and McGowan follow suit. They say their account generates 'the right results' in the classification of things like contracting speech, discriminatory speech, and criminal solicitation, by which they mean that their account matches up with the already widely-accepted position that these kinds of communicative acts shouldn't qualify as SPEECH. But the problem, again, is that it's very hard to extrapolate from this, given that we don't have any independent conception of *significance* (in relation to 'obligation-enactment') to govern our classification. Is a 'significant' obligation enacted when someone greets me on the sidewalk (i.e., an obligation to reciprocate), or when someone writes something silly on the internet (the obligation to correct them)? Naturally Maitra and McGowan want to say that these obligations aren't significant. But how does their account actually substantiate that verdict? They say these obligations are

religions on this issue). I'm just saying that hate-filled carry-on of this kind, laced with threats of eternal punishment, is preached with some regularity at some evangelical churches.

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press

> *Easily met*, and there would be no *serious* repercussions even if those obligations went unmet. By contrast … "You're fired, said to an employee … is significantly obligation-enacting, since it frees the employee from *significant* obligations … and creates *significant* new obligations …. Our thesis is just that significantly obligation-enacting utterances ought to be uncovered.[54]

The words I've italicized in the above passage show the extent of the problem one faces in trying to get determinate guidance from the relevant notion of significance. To say 'it's a significant obligation if it's not *easily* met' is to give content to one flexible criterion by invoking another one. When we look at a pair of communicative act-types for which our classificatory judgments are generally settled—say, scholarly research (SPEECH), versus perjury (non-SPEECH)—we can explain these judgements by saying that the latter enacts 'significant' obligations and the former doesn't. But when we turn to a more controversial case, like HHP, it's rather optimistic to think that our abstract sorting criterion can disinterestedly arbitrate matters to resolve the classificatory controversy. What will typically occur is that each party will begin with some conviction about whether HHP does or doesn't warrant special protection against government restriction, and from this each party will try to retrofit their appraisal—that HHP does/doesn't enact significant obligations—to substantiate the prior judgment. Where a method of this type aims to provide a clear and consistent classificatory system, it ends up misrepresenting classification-by-fiat as if it were a kind of quasi-objective exercise in the categorization of social intercourse.[55]

The only truly unbiased, properly criterial and systematic way of defining non-SPEECH, is by way of a bare linguistic criterion: it's SPEECH if it employs linguistic tokens, otherwise it is non-SPEECH. At least one author opts for something like this, suggesting that "a fundamental difference obtains between symbolic and non-symbolic interactions", and that free speech concerns are "always implicated in the former, and only occasionally in the latter".[56] But this definitional method is pointless. Once we detach our definitional task (i.e., distinguishing SPEECH from non-SPEECH) from our larger theoretical task (i.e., formulating a definition that answers to the ARGUMENTS), we can define SPEECH however we like, since nothing at all hinges on it.

---

54. Maitra & McGowan, *supra* note 44 (2010) at 352 [emphases added].

55. In some respects Maitra and McGowan's work on this topic is an artefact of the dialectic that they're responding to in First Amendment doctrine and discourse. Where Greenawalt favors his approach because he thinks there is something like a natural-kind distinction between SPEECH and non-SPEECH, which simply requires illumination, Maitra and McGowan use his type of methodology primarily as a creative way to show why things like anti-hate speech laws (which they think can be justified in the abstract) can also be made compatible with established First Amendment doctrine, despite prevailing opinion to the contrary. If hate speech is customarily classified as SPEECH, and if this makes it effectively impossible to legally regulate it, then—as a practical argumentative aim—adapting a well-regarded piece of First Amendment theorizing about coverage to argue that hate speech shouldn't be covered by free speech principles is a sensible and savvy maneuver. My criticisms here are naturally only concerned with the deeper theoretical adequacy of this methodological approach to free speech coverage questions.

56. Franklyn S Haiman, *"Speech Acts" and the First Amendment* (Carbondale: Southern Illinois University Press, 1993) at 5.

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press

7. The Subtraction by Comparison Strategy

In trying to carry out the required definitional work, we can't fully avoid the influence of biases, dogmas, and conventional presumptions. The challenge is to find a method for defining SPEECH which, as best as we can manage in spite of this, answers to the ARGUMENTS. I've argued that we get nowhere by seeking a criterial formula that distinguishes SPEECH from non-SPEECH. But the root problem here had better not be the bare fact that categorizations are required. If principled categorization of different kinds of acts is impossible, then Fish was right: free speech theory is hogwash, irrespective of our method for defining SPEECH. So how can we advance? Where's the solid ground to build on? There are some uncontroversial paradigm cases—of both SPEECH and non-SPEECH—that might serve as reference-points in our classificatory task. Anyone with any sympathy at all for free speech agrees that SPEECH should encompass things like scientific research, political dissent, and subversive art. Similarly, practically everyone who believes in a politics of free speech agrees that SPEECH should *not* encompass something like perjury; that is, the fact that people use words when perjuring themselves doesn't mean that anti-perjury statutes should be regulated by a limiting, free-speech-based side-constraint. So here's another kind of Subtractive Approach to defining SPEECH. Identify some forms of communicative action which incontrovertibly (by the lights of all interested parties) don't qualify as SPEECH, and use these as a comparative guide for subtracting any other types of communicative acts whose malign social operations are relevantly similar to these paradigmatic instances of non-SPEECH.

In this type of subtraction by paradigm-case comparison we still need *some* explanation of why the paradigm case is non-SPEECH. But unlike in Greenawalt's method, this explanation needn't involve an appeal to criteria that purportedly define non-SPEECH in abstract, general terms. Take the example of a restaurateur putting a sign in her window that says 'whites only'. The malign operations of such a communicative act are obvious, and could be spelled-out in a number of different directions. Suffice it to say, minimally just societies prohibit discriminatory segregation based on race, and in such societies to hang a 'whites only' sign is *ipso facto* to perform an unlawful discriminatory act. This seems like a good reason to count 'whites only' signs among the communications that don't qualify as SPEECH. But have we provided a general criterion for defining non-SPEECH? No. All we should say is that 'whites only' signs are one of the communications to be subtracted from SPEECH. The considerations that favor this judgment may or may not be related to the reasons why 'whites only' signs are properly criminalizable, but all we're committing to is the view that if other communicative acts serve as tools of segregation and discrimination, like 'whites only' signs do, then they should likewise be subtracted from the domain of SPEECH. This method needn't generalize beyond the specific verdict, e.g., by claiming that being discriminatory or unlawful is what makes communication non-SPEECH.[57]

57. Sunstein appears to endorse the view I'm cautioning against here. He says all communicative action should qualify as SPEECH unless it "amounts to the commission of an independently

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press

This is the type of method that's used in a more recent piece by McGowan,[58] in which she argues that targeted, face-to-face expressions of racist sentiment, in public spaces (e.g., on the subway), operate like 'whites only' signs—that they are, in her words, 'verbal acts of racial discrimination'—and should be subtracted from SPEECH and disentitled to special protections accordingly. The argument that McGowan offers in defense of her claims about the functional equivalence of the two kinds of communicative acts is a plausible one, although it's not directly relevant to the key point I want to make here.[59] The point is about this method for distinguishing SPEECH from non-SPEECH, and it's that if we are committed to following some version of the Subtractive Approach to defining SPEECH, this is about as good as things will get. If we attempt to formulate any general criterion for communicative non-SPEECH, we'll end up indulging in the false pretenses to systematicity and impartiality that were evident under the Greenawalt method. Muddling through, in the space between policy-making and constitutional interpretation—à la First Amendment jurisprudence—will transform the question of how SPEECH is defined into an occasion for culture warfare and conceptual gerrymandering. The best we can do is take whatever cases there are that we all agree should be subtracted from SPEECH—*like* perjury, *like* 'whites only' signs—and by way of piecemeal pair-wise comparisons, see whether any further communicative acts can be singled-out for subtraction.

If this is right—if 'subtraction-by-comparison' is the least problematic version of the Subtractive Approach—and if this method for defining SPEECH is of decidedly limited use, due to its very structure, then it's time to consider wholesale alternatives. As a way into this, consider a strand of thought that is prominent in the free speech tradition right back to Mill,[60] that's in effect in the classic examples

---

illegal act" or "is evidence that the act has been committed"; Cass R Sunstein, "Words, Conduct, Caste" (1993) 60:3 U Chicago L Rev 795 at 836. But this won't do. In a society where blasphemy is criminalized, for instance, the Courts could profess a commitment to Sunstein's brand of free speech, just by noting that blasphemers are engaged in *criminal conduct*, such that no SPEECH is in fact restricted in their being punished. Determinations about which acts are criminalized ought to be answerable to prior judgments about what qualifies as SPEECH, not *vice versa*. Sunstein is right to observe a distinction between cases in which an utterance *constitutes* a criminal act (like a threat), and cases in which an utterance immediately *causes* a criminal act (like an incitement). Still, in some cases we should privilege the former as SPEECH. Yelling "the King is an idiot" on the street doesn't cause sedition, it constitutes sedition. But nevertheless, sedition (in at least some of its forms) is a type of conduct that liberal free speech principles should protect, not something to be placed outside their sphere of protection simply on account of its being identified *as* a type of conduct.

58. Mary Kate McGowan, "On 'Whites Only' Signs and Racist Hate Speech: Verbal Acts of Racial Discrimination" in Ishani Maitra & Mary Kate McGowan, eds, *Speech and Harm: Controversies over Free Speech* (Oxford: Oxford University Press, 2012) 121. This piece is not tightly linked to the co-authored pieces discussed in §§5-6, Maitra & McGowan, *supra* note 44 (2007; 2010), which emulate Greenawalt's program of seeking to identify the criterial features of verbal non-SPEECH.

59. The argument, in brief, is that both communicative acts are 'exercitives'—speech acts that enact permissions/prohibitions—which have the capacity to enact permissibility facts by virtue of general features of pragmatic accommodation in conversation, and which enact permissibility facts that are discriminatory. This account builds on McGowan's earlier work on conversational exercitives as developed in, e.g., Mary Kate McGowan, "Oppressive Speech" (2009) 87:3 Australasian J Phil 389.

60. On one hand, Mill characterizes SPEECH in terms of 'liberty of thought and discussion' and develops ARGUMENTS in defense of it based on that characterization; but at the same time, as

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press

of free speech exceptions from Mill and Holmes (see §1), which is made explicit in Greenawalt's work,[61] and which is woven into McGowan's arguments about 'whites only' signs. 'Whites only' signs don't qualify as SPEECH, McGowan says, "exactly because of what they *do*", and because "we have reason to believe that some racist hate speech *does* the very same thing, we have reason to place that subset of racist hate speech outside the scope of a free speech principle".[62] What I want to flag here is the thesis that's implicated when we stress what a communicative act *does* in order to explain why it shouldn't qualify as SPEECH. The implicature of this is that communicative acts qualify as SPEECH, except in those instances in which somebody's words burst out of the inert realm of ideas, and transmogrify into 'actions', 'doings', or social 'moves'.[63] McGowan quite rightly insists that words *aren't* always inert and ethereal; that they *do* 'change stuff'. But the implication—here and also in many quarters of free speech theory—is that it's precisely in those cases in which words 'change stuff' that we should subtract them from the domain SPEECH. However, the communicative acts we're trying to protect from government overreach, in a political order regulated by free speech principles, do not exist in a quasi-inert realm of pure ideas. Much of what we want to specially protect is communicative action that impacts upon and changes the world, some of it for better, some of it—in the short term, at least—for worse. The Subtractive Approach coaxes us into ways of conceptualizing the SPEECH v. non-SPEECH distinction which locate all capacity for impact and change on the non-SPEECH side of the divide. And that's a distortion—a commonplace distortion, but a distortion nonetheless—of the political concerns that underwrite free speech ARGUMENTS.

8. Addition instead of Subtraction?

One step towards ridding ourselves of these distortions may be to take note of the classificatory method that was adjacent to the one McGowan offers. There are indeed paradigm cases of *non*-SPEECH, like perjury and 'whites only' signs, whose

Haworth says, there is "a great deal which [Mill] does want to allow, which doesn't easily match the paradigm, but which he nevertheless wants to defend in terms of the same argument he applies to thought and discussion"; Alan Haworth, *Free Speech* (London: Routledge, 1998) at 32. Mill fails to see how much diversity there is in the kind of communicative action that the liberal state might want to protect, Haworth says, because under his analysis, everything that's worthy of protection must be shoehorned into the conceptual category of 'thought and discussion'.

61. For example, in the following: "A pervasive theme of this book is the difference between saying something with words and doing something with words. I have stressed the importance of the speaker's aims in deciding how particular speech should be regarded. But many of the subjects we have considered show how much speech importantly does something, apart from influencing actions because people are persuaded by the ideas that are communicated. Many of those who claim that speech … undermines equality emphasize what this speech does, and they call on the government to stop the harmful effects"; Kent Greenawalt, *Fighting Words: Individuals, Communities, and Liberties of Speech* (Princeton: Princeton University Press, 1995) at 152.

62. McGowan, *supra* note 58 at 145 [emphases added].

63. Here I'm echoing the kinds of expressions used by critics of free speech orthodoxy like Fish, *supra* note 35 at 125-26, and MacKinnon; see Catharine A MacKinnon, *Only Words* (London: Harper Collins, 1994) at 8.

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press

status as non-SPEECH is uncontroversial among proponents of free speech. But there are also uncontroversial paradigm cases of SPEECH that may likewise serve as informative reference-points for our classificatory task. As I suggested above, all free speech sympathizers will agree that SPEECH should encompass academic research, political dissent, and subversive art. What makes these things paradigm cases of SPEECH, in short, is that they so clearly implicate the benefits and/or dangers adverted to in the ARGUMENTS. Instead of trying to differentiate SPEECH from non-SPEECH by extrapolating (by comparison) from paradigm instances of non-SPEECH, then, couldn't we construct our classificatory system by extrapolating (by comparison) from these paradigm cases of SPEECH itself? This method we're edging towards is an inversion of the Subtractive Approach, as foreshadowed in my title, i.e., the 'Additive Approach'. Schauer envisioned all this in an old paper, describing the two methods as 'defining-out' and 'defining-in'.[64] Rather than categorizing all acts of communication as SPEECH by default, then trying to make principled subtractions to exclude the communicative acts that don't merit the special protections owed to SPEECH, we could instead refrain from categorizing anything as SPEECH by default, and only install things in that class via principled *additions* which *include* the communicative acts that *do* merit the special protections owed to SPEECH.

If a straightforward endorsement of this Additive Approach was the intended upshot of my critical analysis in §§3-7, there would be at least two objections in the offing. Here's the first one. These two approaches, Subtractive and Additive, are both methods for categorizing acts based on whether, in light of the ARGUMENTS, those acts do or don't merit special protection against legal restriction. If the conditions under which these methods are applied allow for clear-sighted readings of the ARGUMENTS, and of which act-types are implicated by them, then the two methods should yield *identical* categorizations. Assuming that our conditions are not conducive to such perfect perspicacity, the difference between the two methods will, in practice, just be a difference in types of error. Errors under the Additive Approach will be false negatives; things that should be protected under the auspices of free speech won't be. Errors under the Subtractive Approach will be false positives; things that shouldn't be protected under the auspices of free speech will be. Schauer thinks this counts against the Additive Approach, because we should "attempt to ensure that decisions under uncertainty will be biased away from the restriction of those values we hold to be of greatest importance".[65] In other words, overprotection is better than under-protection, given that the goods we're promoting and dangers we're guarding against are especially weighty. But the Libertarian slant that guides Schauer's judgment here is ultimately tenuous. Free speech discourse does well when it comes to doomsaying about the dangers of under-protection, of governments not being subject to robust enough constraints on their attempts to legally suppress communicative conduct. These dangers are real, and vigilantly guarding against them is the right response. But free

64. Frederick Schauer, "Categories and the First Amendment: A Play in Three Acts" (1981) 34:2 Vand L Rev 265.
65. *Ibid* at 281.

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press

speech discourse does a less impressive job when it comes to doomsaying about the genuine and countervailing dangers of *over*protection. If dual-use research, say, or genocidal propaganda, are accorded too robust an immunity against legal restriction, then the dangers of this are—I'll err on the side of understatement here—*at least as bad* as the perils emphasized by theorists who doomsay about under-protection. To deny this is either to exaggerate the badness of censorious tyranny—by pretending that it's worse than a global pandemic, or mass genocidal violence—or simply to play dumb about the issues.[66] The proper thing to say, is that there are fateful dangers associated with overprotection and under-protection alike. If the likelihood of false negatives on the Additive Approach is equivalent to the likelihood of false positives on the Subtractive Approach, then in relation to this criterion of assessment, we should score the two methods as a draw. What *would* constitute a tie-breaker would be some reason to think that, when put to use in real-world (i.e., in imperfectly-perspicuous) conditions, one method will yield fewer classificatory errors *period*.

But this brings us to a more telling concern about the Additive Approach. There's no real reason to think that switching from subtraction to addition will make it easier to overcome the distortions, biases, and other sources of error that make it hard to discern which communicative acts merit special protection as SPEECH. Whether we're adding or subtracting, the interpretative tasks are basically the same: review the reasons that obtain, in light of free speech ARGUMENTS, for protecting conduct of type α, then examine the communicative act-type that we're trying to classify as SPEECH or non-SPEECH, and see whether it matches the characteristics of α. And if the interpretative tasks are the same under either approach, then the frequency of classificatory errors probably will be too. Switching to the Additive Approach, while leaving all else in place, won't remedy the classificatory difficulties that arise under the Subtractive Approach.

If there's a good reason to prefer the Additive to the Subtractive Approach, then it has to be something like the thought raised at the end of §7, i.e., that the Subtractive Approach coaxes us into thinking about the SPEECH v. conduct distinction in a way that distorts our interpretations of the ARGUMENTS and the communicative act-types they implicate. This is what my preference for the Additive Approach boils down to, and there are two distortions that I'll discuss. The first consists in the notion that the communicative acts encompassed *within* SPEECH are, of their essential nature, forms of non-action or non-conduct. The second consists in us thinking of either classification within that binary as a group whose members all share an essential property or exhibit some common trait. I'll discuss these in §9 and §10 respectively. To recapitulate my point here: if we adopt the Additive Approach to defining SPEECH, while continuing to conceive of SPEECH under the influence of these twin misapprehensions, then we'll

66. On incitement to genocide, see, e.g., Susan Benesch, "Vile Crime or Inalienable Right: Defining Incitement to Genocide" (2008) 48:3 Va J Intl L 485; Lynne Tirrell, "Genocidal Language Games" in Ishani Maitra & Mary Kate McGowan, eds, *Speech & Harm: Controversies over Free Speech* (Oxford: Oxford University Press, 2012) 174. On dual-use research, see, e.g., Nicholas G Evans, "Dual-use Decision Making: Relational and Positional Issues" (2014) 32:3 Monash Bioethics Review 268.

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press

be no better-placed to perform our classificatory tasks than we were in using the Subtractive Approach. Switching to addition only constitutes progress to the extent that it helps us to appropriately reconceptualize the categories of protected communicative acts into which things are going to be 'added'.

9. Speech as Conduct's Antithesis

In some cases the characterization of Speech as communication which is of its *essence* 'not-conduct' is more explicit, like when Charles Collier speaks about the "obviously experienced differences between speech and conduct",[67] or when Martin Redish suggests that it's "almost certainly true in the overwhelming majority of cases that speech is less immediately dangerous than conduct".[68,69] As we observed in §7, however, in the discussion of McGowan on 'whites only' signs, this view of Speech's essential nature is commonly manifested implicitly, when an author is framing the case for subtracting some class of communications from Speech, by stressing the fact that the communications in question "are acts and *not just words*".[70]

The inadequacy of this characterization is evident when we examine the substance of the Arguments and observe the types of communications that they advert to as being in need of special protection. What we see, is that 'Speech' encompasses all sorts of tangible, 'impactful', 'difference-making' forms of behavior—things that are *conduct* plain and simple. (i) A scientist wants to share data with collaborators and publicize her findings. The government wants to impose strict regulatory oversights on this, because it's worried about the costs of dangerous data falling into the hands of its enemies, or lucrative data falling into the hands of parties that didn't invest in it. *Epistemic* Arguments tell us we mustn't allow governments to enact such oversight, based on bare cost-benefit analyses, because they're subject to pressures and self-interested biases in ways that will systematically push them towards overregulation, to our collective detriment over time. (ii) A political radical is circulating messages, telling others to engage in militant political opposition to the Corporatocracy that's destroying the world. The government wants to quell these messages, so as to inhibit opposition to its Corporatocratic elements and shore up the status quo. *Democratic* Arguments tell us we must disallow the state's suppression of troublemaking dissent, again, because of various pressures and biases that systematically

67. Charles W Collier, "Hate Speech and the Mind-body Problem: a Critique of Postmodern Censorship Theory" (2001) 7:2 Leg Theory 203 at 233.

68. Martin H Redish, *Freedom of Expression: A Critical Analysis* (Charlottesville: Michie, 1984) at 5.

69. For persuasive critiques of the notion hinted at in these remarks, i.e., that Speech is of its nature harmless or inherently less harmful that *bona fide* conduct (also sometimes called the 'sticks and stones' view of Speech), see Frederick Schauer, "The Phenomenology of Speech and Harm" (1993) 103:4 Ethics 635; Susan J Brison, "Speech, Harm, and the Mind-body Problem in First Amendment Jurisprudence" (1998) 4:1 Leg Theory 39.

70. Phrases like this crop up frequently in articles on the boundaries of free speech; this phrase comes from David Archard, "Insults, Free Speech, and Offensiveness" (2014) 31:2 Journal of Applied Philosophy 127 at 139 [emphasis added].

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press

motivate governments towards over-suppression, to society's long-term detriment. (iii) An iconoclastic artist plans an exhibition of works which deeply offend certain religious elements. The government wants to shut down the exhibition, to protect those religious elements against perceived religious vilification. *Expressive* ARGUMENTS tell us that we mustn't allow governments to exercise such censorious powers, as this breeds a culture of narrow-mindedness and conformity, to our collective detriment in the long-run. We could multiply examples all day. All of these communications are things that do—and are *intended* to, and are threatened with restriction precisely *because* they—make a potent causal impact on our society.

Others in the literature have occasionally criticized this regrettable propensity to characterize SPEECH as being, of its essence, unlike conduct. In reply to Thomas Emerson's sanguine proposal, that we could classify communicative acts as SPEECH or conduct depending on whether they 'partake' of the essential qualities of expression or action,[71] Edwin Baker hits the mark nicely. "Neither common sense nor the purposes of the system of freedom of expression" he says, "work to distinguish between the 'essential qualities' of expression and action"; rather, put simply, "most conduct falls into both categories".[72] Or to take an earlier example, consider this passage from John Hart Ely's classic paper on flag desecration.

> Burning a draft card to express opposition to the draft is … 100% action and 100% expression. It involves no conduct that is not at the same time communication, and no communication that does not result from conduct. Attempts to determine which element predominates … inevitably degenerate into question-begging judgments about whether the activity should be protected.[73]

So why, if the literature has alerted us to its faults, is there still this impulse to think of SPEECH and conduct as opposites? Often, I suspect, this isn't borne of any actual conviction as to the inertness of 'mere' opinions or words, but just the fact that SPEECH—whether or not it genuinely *is* special—nevertheless *needs* to be special in order for the free speech tradition to make sense (here I'm paraphrasing Schauer).[74] By which I mean: the way SPEECH is conceptualized may be affected by us feeling a need to rationalize its 'specialness', by attributing a special form of inertness to it. In any case, there's only one real argument that I know of, for this view, which is that (some or all) paradigmatic instances of SPEECH—scientific research, protest, iconoclastic art—are things that don't have their impact on the world directly. When communications like these change the world, they do so by affecting the thoughts or beliefs of others, who subsequently alter their own conduct. Of course, affecting people's minds *is* a way of having an effect on the world. Ultimately this argument trades on ideas about responsibility. When A

71. Thomas I Emerson, *The System of Freedom of Expression* (New York: Random House Trade, 1970) at 18.
72. Baker, *supra* note 16 at 70, 71.
73. John Hart Ely, "Flag Desecration: a Case Study in the Roles of Categorization and Balancing in First Amendment Analysis" (1975) 88:7 Harv Law Rev 1482 at 1495.
74. Schauer, *supra* note 4 at 1306.

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press

says to B 'anthropogenic climate change is all made up', and B then decides to trade-in her Prius to buy a Hummer, we attribute responsibility for this choice to B. We recognize that A's SPEECH causally instigates B's purchase, but we judge that the outcome results from B's agential intervention, in a way that 'screens off' A's accountability for it. This type of thinking about what SPEECH is, and why it should be specially protected, figures in the work of plenty of free speech theorists,[75] and issues about responsibility for SPEECH's effects are naturally to the fore in work on incitement to violence.[76] Even if there isn't any neat and natural distinction between words that alter the world indirectly (i.e., SPEECH) and those that alter the world directly (i.e., non-SPEECH or communicative conduct), it's hard to deny that there is something important in this line of thought. There is some axis of difference here, and the concepts that we reach for to describe this difference—autonomy, persuasion, intellectual agency—are thematically central to the force of free speech ARGUMENTS. The question is how to acknowledge this, without recommitting the error of imagining that the things the ARGUMENTS tell us to protect, as SPEECH, exist in an inert domain of mere ideas. One option would be to keep invoking the SPEECH v. conduct distinction, much the same as ever, and just make a more concerted effort to resist the mischaracterizations that I've been criticizing. No more saying 'α shouldn't be protected under the auspices of free speech, since it's more of an *act* than an *idea*', thus implying that the things to be protected by free speech *aren't* actions, or are essentially unlike actions. Instead, we might acknowledge that various paradigmatic instances of SPEECH—protest, scholarship, art, etc.—achieve their effects indirectly, by affecting the thoughts of other people, while simultaneously emphasizing that they're *bona fide*, world-altering actions. We could attempt to stress the 'causal potency' of the communications that we want to specially protect, as SPEECH,[77] while asserting the necessity of their protection all the same.

The proposal I want to advance, however, represents a more dramatic departure from theoretical orthodoxy. Having spent a good while diagnosing the difficulties that come with defining 'SPEECH' in a way that answers to those ARGUMENTS that confer meaning on that definitional task, I'm now going to argue that our talk of 'SPEECH', as an overarching category denoting the things that merit special protection, is in fact part of what compromises definitional clarity in this arena. 'SPEECH' should ultimately be subdivided, and then jettisoned.

---

75. See Thomas Scanlon, "A Theory of Freedom of Expression" (1972) 1:2 Philosophy & Public Affairs 204; Dworkin, *supra* note 15; Husak, *supra* note 4; David A Strauss, "Persuasion, Autonomy, and Freedom of Expression" (1991) 91:2 Colum L Rev 334; Thomas Nagel, "Personal Rights and Public Space" (1995) 24:2 Philosophy & Public Affairs 83; Evan Simpson, "Responsibilities for Hateful Speech" (2006) 12:2 Leg Theory 157.

76. See Lawrence, *supra* note 24; Larry Alexander, "Incitement and Freedom of Speech" in David Kretzmer & Francine Kershman Hazan, eds, *Freedom of Speech and Incitement Against Democracy* (London: Kluwer Law International, 2000) 101.

77. As in passages like this: "Marx and even Bentham have probably nurtured more acts of terrorism than all the soapbox firebrands put together …. Does that mean that [governments] should ban Marxist and consequentialist philosophy or impose a religious orthodoxy on its citizens? To ask that question is to answer it, for any country that pursues security through such repression is … unworthy of salvaging"; Alexander, *supra* note 76 at 118.

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press

10. Eliminating Speech by Subdivision

Free speech Arguments run along at least four distinct trajectories. And while the cumulative force of the Arguments is usually taken to be that we should accord a special status to a certain class of communicative acts, i.e., Speech, there is no unified—certainly no simple and intuitive—class of acts that the Arguments nominate to receive that privileged status. This is a paraphrase of a passage that Thomas Scanlon wrote more than 40 years ago.[78] Thoughts of a similar kind also crop up repeatedly in Schauer's work, such as in the following passage.

> Freedom of speech need not have any one 'essential' feature. It is much more likely a bundle of interrelated principles sharing no common set of necessary and sufficient defining characteristics … the protection of political discussion and criticism, the aversion to censorship of art, and the desire to retain open inquiry in science and other academic disciplines … are principles not reducible to any one common core. Any attempt to do so is likely to … distort all of the principles involved.[79]

The idea that 'free speech' refers to a heterogeneous grab-bag of communications, and not some unified class of communications, isn't new. My proposal is to take this idea more seriously, and restructure the categories that organize free speech principles accordingly. Free speech can be divided into a plurality of principles, each drawing from a particular type of Argument, and imposing its own constraints on government action. In place of a 'free speech principle' *per se*, this would give us several more narrowly-focused principles, which jointly provide protections for all the types of communicative action which really do merit it, in light of the Arguments. So: protest, satire, and political argument enjoy special protection against legal restriction, being selected for this privilege by a *Freedom of Protest/Politicking* principle. Journalism, scientific inquiry, and academic debate enjoy protection against legal restriction, being nominated for this by a *Freedom of Inquiry* principle. Art, literature, theatre, comedy, and cinema enjoy special protection against legal restriction, being nominated for this by a *Freedom of Art* principle. And proselytism, religious texts, and ethical and spiritual advocacy enjoy special protection against legal restriction, being selected for this privilege by a *Freedom of Conscientious Expression* principle.

The following table maps out these schematic, three-part relationships.

78. Scanlon, *supra* note 75 at 208.
79. Schauer, *supra* note 64 at 277.

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press

Arguments	Communication Types	Principle
Democratic Arguments: failing to specially protect certain kinds of communications undermines the legitimacy and/or effectiveness of democratic politics.	Protest, political dissent, criticism of public figures, satire, political debate	Freedom of Protest/ Politicking
Epistemic Arguments: failing to specially protect certain kinds of communications threatens various important social epistemic goods.	Scholarly research, journalism, reportage, public debate, informal and private research, other inquiry	Freedom of Inquiry
Expressive Arguments: failing to specially protect certain kinds of communications jeopardizes goods related to creativity, authentic cultural expression, and pleasure in the arts.	Literature, visual arts, comedy, cinema, music, theatre, dance, avant garde performance	Freedom of Art
Worldview Arguments: failing to specially protect certain kinds of communications tends to lead to evils related to enforced orthodoxy, dogmatism, and the suppression of heresy.	Proselytism, circulation of religious or ethical texts, advocacy on behalf of a particular cultural, spiritual, or ethical way of life	Freedom of Conscientious Expression

This approach isn't committed to treating the four areas that I've identified here as the final word in free speech Arguments. Some theorists will think other Arguments should be added into the mix (see note 19). Some will think the Arguments that I've presented should be merged, or explicated differently.[80] What I'm ultimately endorsing is just this kind of schematic approach, which would obviate the need for a definition of Speech *per se*, because it would abandon the pretense that the cogent free speech Arguments—if not these four, then the Arguments that *are* cogent—recommend special protection for some unified family of behaviors. Instead, it looks at each free speech Argument, ascertains the distinctive types of communication therein singled-out for special protection, and postulates a separate side-constraint on government action which accords special protection to just those particular species of communication.

To properly expound a free speech theory of this kind, much work is required beyond this inaugural restructuring maneuver. If we are applying a plurality of

80. For instance, among the three kinds of fundamental interests which, on Cohen's view, are secured by a system of expressive liberty, one of them ('informational interests') roughly corresponds with the values adverted to in Epistemic Arguments, and another of them ('expressive interests') corresponds very closely with the values adverted to in Expressive Arguments. But the third kind of fundamental interests ('deliberative interests'), implicates values that, on my taxonomy, are adverted to by two different kinds of arguments, namely, Democratic and Worldview Arguments; see Cohen, *supra* note 13 at 223-29.

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press

distinct side-constraints on government action, instead of a single multifaceted side-constraint (as in standard free speech theory), we'll have to determine how each side-constraint operates by itself, and how they all interact. Some communicative entities—like artworks that express political dissent—will be covered by multiple principles, and thus questions about precedence of coverage will have to be addressed. Some of the newly-minted principles will demand something like standards of strict scrutiny, whereas others may only require intermediate scrutiny, or some other standard of legislative review. And the way that we think about content-neutrality could vary between the different principles; in other words, what counts as a content-based restriction may need to be assessed differently under, say, Freedom of Art versus Freedom of Inquiry. In short, the intricate machinery of legal policy will still have to be painstakingly engineered, in these and other respects. However, that's not a reason for proponents of a standard (non-subdivided) approach to free speech theory to dismiss what I'm proposing. *Any* sophisticated system of free speech that purports to answer to the ARGUMENTS will posit differential levels of protection, sometimes independent, sometimes overlapping, for different species of communicative acts.[81]

The real question is what we gain by jettisoning the umbrella category of SPEECH, and restructuring principles of communicative liberty in this pluralistic, ramified fashion. Throughout my discussion, I've emphasized 'answering to the ARGUMENTS' as the most important attribute of a sound definitional method in this arena. If we're going to specially protect a certain class of acts against legal restriction, then we need to describe and delimit the protected class in a way that doesn't readily allow acts to be accorded this privilege unless they're actually identified, by the ARGUMENTS, as meriting it. None of the ARGUMENTS can plausibly be read as entailing that special protections are needed for something which would be fittingly denoted using a term as wide-reaching—and as susceptible to equivocal construals—as 'SPEECH'. And therefore, by treating the ARGUMENTS as if they do all enjoin special protection for one class of behaviors, i.e., some multifarious set of communications denoted by 'SPEECH', standard free speech theory flunks the 'answering' test; it makes it too easy for acts to be accorded special legal privileges, even when they bear only a tenuous connection to the social processes via which the goods and ideals adverted to in the ARGUMENTS are advanced or endangered. In saying this I don't mean to claim that in a subdivided free speech system it would *impossible* for acts that aren't nominated for protection by the ARGUMENTS to be accorded it anyway. The institutions that implement our principles of communicative liberty will still be staffed by (biased, error-prone) human beings, and they'll still be situated in (factionalized, corruptible) political communities. My contention is just that we'll do better at protecting only those acts that genuinely merit it, if our classificatory questions and answers refer to narrower and more specific categories of communicative

81. Another issue with this approach is that it may turn out, on close inspection, that the different ARGUMENTS have culturally specific boundaries and culturally specific applications which cannot be adequately captured at the level of abstract argument and principle; on this point see Koppelman, *supra* note 42 at 700.

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press

conduct, i.e., if we're asking whether α is entitled to special protection as an instance of PROTEST, instead of asking whether α counts as SPEECH.[82] Or to put it another way: definitions of protected classes of communication will do better at answering to the ARGUMENTS, if the protected classes that they're defining—like PROTEST, or JOURNALISM, or RELIGIOUS TEXTS—are things whose nature and boundaries can be discerned from the substance of the ARGUMENTS themselves. In standard free speech theorizing, the class of protected communicative behavior that needs to be defined, SPEECH, is an abstraction that's at one remove from the substance of any of the ARGUMENTS, insofar as it's meant to reach across them all, and encompass a unified class of communications to which they all supposedly refer. The difficulty of defining this class in a way that answers to the ARGUMENTS isn't so much the fault of any definitional method, as it is a result of us asking one umbrella concept to do more work than it realistically could ever have handled.

My focus has shifted in this section, from questions about *how* acts get sorted into the categories that a free speech theory defines, to the question of *what* categories a free speech theory should seek to define in the first place. It may seem that the Subtractive Approach and its alleged defects have dropped out of the analysis. But that's not quite right. Consider an analogy. Suppose we're trying to map out a binary taxonomy, distinguishing the living things that are ANIMALS from the ones that aren't. Given this kind of binary sorting task, it makes good sense to ask whether an Additive or Subtractive method is going to do a better job at correctly classifying the entities in question. But if we're trying to taxonomically sort living things into multiple classes—FISH, MAMMALS, INSECTS, BIRDS, etc.—then an Additive definitional method would clearly strike us as the better approach from the outset. We *could*, as per the Subtractive Approach, classify all living things as FISH by default, and then specify criteria that identify living things as non-FISH, so as to subtract them from that class, before repeating this method for each category in turn. But that would be an absurdly convoluted way to carry out the task. The more efficient and reliable method would obviously just be to specify what makes something a FISH, and then add things into that taxonomical class if they satisfy these criteria. A Subtractive Approach to a definitional problem only looks like a viable methodology when the definitional problem is a binary sorting task. The point, then, is that if we favor—or uncritically accept—a Subtractive Approach to defining categories of specially protected communicative conduct, we are prematurely rejecting a Subdividing Approach to free speech theory. In short, an account of the problems that come

82. To see that something like PROTEST is not merely a relabeled version of our original SPEECH category, we need only note how dissimilar to 'speech' much of what merits protection *qua* PROTEST is. Setting fire to an object—a flag, an effigy, or a bra—can all be acts of protest. Conspicuous and defiant silence—that is *literal* silence: refusing to speak when called upon to do so—can function as protest under certain circumstances; see Louis Michael Seidman, *Silence and Freedom* (Stanford: Stanford University Press, 2007). All of these acts should be accorded special protection against coercive government interference. But this is not because they are 'SPEECH'. Rather, it's because they are instances of a type of activity, PROTEST, which may or may not be performed via the use of linguistic tokens, but which, so the Democratic ARGUMENTS indicate, we have good reasons to specially immunize against government restriction irrespective of their linguistic or non-linguistic character.

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press

with a Subtractive Approach to defining SPEECH is not merely a way of setting the stage for consideration of an unorthodox, subdivided conception of free speech theory; it's more like a necessary prerequisite for our envisioning this alternative.

11. The Alleged Prudence of Casting a Wide Net

In §10 I said nothing about ARGUMENTS from Government, which treat the incompetence or mendacity of governments as a key justificatory basis for free speech. Rather than training our focus on any particular species of communicative action (e.g., the way *Epistemic* ARGUMENTS train our focus on inquiry-oriented communications, like scholarly research), ARGUMENTS from Government express general skepticism about the capacity of governments to regulate *any* communicative conduct in a way that competently and impartially furthers only legitimate ends. Such ARGUMENTS are grounded in what Joshua Cohen calls 'the Unhappy Facts of Life', chief among which is that people who acquire political power tend to be (and/or become) vain, biased, corrupt, prone to inflating the dangers of the ideas they object to, and so on.[83] Cohen's worries are eminently reasonable, and moreover, they are unmistakably an important thematic touchstone in the free speech tradition. Note, however, that a generalized mistrust of government, in view of the Unhappy Facts of Life, is also a crucial part of what recommends liberalism *per se*, with its aversion to states being governed by a substantive conception of the good, over the more overtly moralized (and thus, more corruptibly parochial) political views on offer in the Communitarian region of the political landscape. If the Unhappy Facts are to underwrite any distinct principles of *communicative* liberty, beyond axiomatic liberal precepts (like the non-regulation of harmless or self-affecting action), then we need some reason to think that these Unhappy Facts are liable to produce more error, or greater corruption, in a government's attempts to restrict certain specific communicative activities, than in its attempts to restrict other kinds of behavior.[84]

In short, there is a burden of specification, for those whose free speech theory supposedly rests on ARGUMENTS from Government, to say precisely *which* communicative acts occasion an additional level of disquiet about government incompetence or mendacity, so as to activate the supplementary normative apparatus of 'free speech'. And this brings us right back to square one. The claim cannot just be that the restriction of 'speech', in the everyday sense, is what triggers a heightened level of disquiet, because there are so many kinds of 'speech' that

83. Cohen, *supra* note 13 at 233.

84. A different strategy for trying to mitigate the perils of the Unhappy Facts would be to focus on constraining government's power to enact legislation of any kind—regardless of whether it suppresses SPEECH—if and when that legislation's purpose is the suppression of disapproved viewpoints; see Eric Barendt, *Freedom of Speech*, 2nd ed (Oxford: Oxford University Press, 2005) at 78-83. But this doesn't obviate the need to specify the acts for which special communicative liberties apply, since there will still be a range of cases in which, even if the government's overt purpose isn't to suppress a particular viewpoint, its actions will nevertheless have that effect, and where we thus need to decide whether what's being suppressed falls under a principle of communicative liberty, such that the government action has to satisfy a more stringent justificatory standard.

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press

evidently bear no relation to the types of communicative acts that are adverted to, in the ARGUMENTS, as exemplars or embodiments of distinctive benefits or dangers. If ARGUMENTS from *Government* tell us anything about the class of behaviors whose restriction activates this supplementary normative apparatus, it is only that some such class exists and demands our attention. In order to progress beyond this point, we need to do precisely what I've been proposing: analyze the ARGUMENTS in the free speech tradition in order to identify the types of communicative acts they nominate for special treatment. What we see, then, is that there are several distinct types of communicative acts that are at issue, not some capacious umbrella class of things called 'SPEECH'.

Someone might argue that we need to retain the umbrella class of SPEECH as a diagnostic category in free speech theory, by again expressing anxiety about the perils of under-protection. Greenawalt for one gestures in this direction. If we doubt our ability to accurately nominate the types of communicative action that require special protection, he suggests, then we actually have reason to welcome the fuzzy boundaries that come with talk of SPEECH *per se*, since they serve to cast a wide net of protection over communications of all varieties. Greenawalt says that

> Sensitive legislators could believe that even their ability to perceive the reach of the justifications for free speech is tainted and that they should accord protection to forms of communication even when they are dubious that the justifications for protection are relevant. This view would be reinforced if the legislators thought that a mistaken suppression of speech was a much worse wrong than a failure to adopt a desirable prohibition of communicative activities.[85]

There are three things to say on behalf of a Subdividing Approach to free speech in reply to this line of defense for an overarching SPEECH classification. (i) This defense ultimately concedes the main point of contention. Using an over-general, equivocally-labelled classification (i.e., SPEECH) to structure the policy system that protects special communicative liberties, ensures that some communicative acts that don't merit special protection are accorded it nonetheless. Against this bias in favor of overprotection, a Subdividing Approach seeks neither to overprotect nor under-protect, but rather to protect all *and only* the communicative acts that *really do*, by the lights of the ARGUMENTS, merit special protection. (ii) To favor overprotection for communication in spite of this, is to engage in a kind of Libertarian 'double-counting', the upshot of which is that governments will be obstructed in their efforts to police some genuinely harmful communicative acts, acts which aren't identified in any ARGUMENT as meriting the protections that underpin the obstruction, but which are protected solely because they get slotted into a nebulous classification and caught up in the wide net that it casts. As I said in §8, the theorist who, in principle, prefers deliberate overprotection to attempts at accurate protection is either ignoring the real dangers of overprotection, or exaggerating the perils of under-protection. (iii) It may be that the integrity of a legal regime at a particular place and time becomes so badly degraded that, in

85. Greenawalt, *supra* note 43 at 41.

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press

order to guard against unchecked tyranny, it's prudent to beef-up our constraints on the government's ability to restrict people's conduct. But if this is what justifies SPEECH's wide net, then there's no reason to limit ourselves to overprotecting communicative acts as distinct from other kinds of activity. Once we abandon the project of seeking to specially protect all and only those things that really do merit special protection, and instead aim to specially protect all the things that actually merit special protection, *plus* some fuzzily-defined penumbra of other superficially similar things, no principled reasons can be invoked in order to set boundaries around that penumbra.

I don't mean to deny the appropriateness of the pessimism about government action that animates ARGUMENTS from Government. One need only revisit the U.S.'s "long and unfortunate history of overreacting to the dangers of wartime", to see how rapidly a regime of expressive liberty can crumble—even a regime which expressly glorifies this liberty—when placed under significant pressure.[86] However, it is precisely because we want to formulate political systems which guard against these Unhappy Facts that we situate our theorizing in the liberal tradition in the first place, and recognize the need for further protections, beyond those already baked into a liberal harm principle, for certain kinds of communicative acts. It is clumsy and artificial (not to mention neurotic) to attempt to structurally encode Libertarian biases into every classificatory element of the principles we're formulating. If a theorist were really *that* pessimistic about the government's ability to implement a principled legal framework as its architects envisioned, then regime-designing projects would be futile by that theorist's own lights anyway.[87]

12. Freedom of Speech Acts

How do the definitional and categorizational issues that I've been discussing connect up with larger questions in free speech theory? There are plenty who dislike what they see when they look at free speech theory, because of how typical understandings of free speech allow that certain racist or sexist practices can qualify as SPEECH, thus forcing arguments for their legal regulation to satisfy more onerous justificatory standards. Some theorists, like Fish (see §4), present such criticisms on behalf of a wholesale rejection of free speech theory. Others—like Maitra and McGowan (see §§5-6)—attempt to redraw the boundaries of SPEECH in a way which shows that, on principled grounds, some of the problematic communications usually taken to sit inside SPEECH's protective sphere can in fact be relocated outside of it. Both strategies aim, in different ways, to chisel

86. Geoffrey R Stone, *Perilous Times: Free Speech in Wartime; From the Sedition Act of 1798 to the War on Terrorism* (New York: WW Norton & Company, 2004).

87. This isn't to say that liberals should just give up on trying to constrain tyranny, but rather to stress that *de jure* legal safeguards ultimately can't achieve this alone. As Stone says, to resist a slide towards tyranny during a crisis "a nation needs not only *legal* protection of civil liberties but a *culture* of civil liberties … an environment in which citizens are more informed, open-minded, skeptical, critical of their political leaders, tolerant of dissent, and protective of the freedom of all individuals". *Ibid* at 537.

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press

away at the aggrandized status that's ascribed to free speech by many who write about it, especially those First Amendment theorists who are eager to wax pious about American Constitutional Exceptionalism. I have reservations about this chiseling, however, because I believe the 'Exceptionalists' are at least half-right. The understanding of free speech that's prevalent in U.S. Constitutional jurisprudence has a radical and progressive idea at its center: that constituencies who attain political power should ostensibly be *banned* from policing the key communicative activities of their opponents, and the relevant forms of communication should be accorded an immunity from government interference that is in some genuine sense *categorical*. In view of the Unhappy Facts noted in §11—the corrupting influence that political power has upon those who wield it, the gross immiserations inflicted by the censorship regimes that have arisen in many political communities—this categorical approach to free speech has the potential, so the Exceptionalists say (rightly, I believe), to underwrite a genuinely emancipatory politics. But this radical and progressive idea generates problems of its own if it is manipulated so as to extend special immunities to types of communication that have no genuine claim to this privilege, and which are merely swept into an overbroad protective net. Is there a way to remain vigilant in protecting the communication that really needs to be kept safe against the repressive urges of the politically powerful—research, journalism, political debate, satire, protest, ethical and religious discourse, art, and literature—without unduly fettering the state's power to enact its legitimate purposes in regulating harmful instances of other forms of communication that *don't* merit the same privileged status? My project here—overturning the Subtractive Approach to defining SPEECH, and offering a subdivided reconceptualization of our categories of special communicative liberty—is an attempt to move free speech theory in this direction.

When one attempts to subtract things like hate speech from the domain of SPEECH, while operating within a conventionally-structured free speech system, the tendency is to reinscribe the longstanding misapprehension, that the things meriting protection as SPEECH are mere *sayings*, whereas verbal acts that don't merit protection as SPEECH are more like *doings*. This warps the substance of the ARGUMENTS that supply free speech theory with its underlying normative force. Democratic ARGUMENTS, for instance, when properly elucidated, don't merely call for special protection for 'expressions of opinion', they demand special protections for (among other things) militant political activism, disruptive satire, the setting-up of election placards on one's lawn, and various other kinds of *actions*. One further thing that bears emphasizing, then, vis-à-vis the motivations for the approach I've been developing, is the way that it chimes with one aspect of the revisionist free speech theorizing we find in contemporary feminist critiques of pornography. In defending Catharine MacKinnon's provocative claim, that pornography silences women, Rae Langton and others argue that pornography effects a kind of 'illocutionary disablement', which is to say that it impairs the ability of women to perform certain illocutionary acts (i.e., to do things with words, like refusing a sexual advance) that they should be able to do, and otherwise *would* be able to do, but for pornography's corrosive influence on the

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press

communicative environment. If this is right, there's at least a *prima facie* case to be made for legally restricting pornography; the state should (legally) silence the speech of pornographers, as a way of preventing them from (illocutionarily) silencing the speech of women attempting to verbally refuse a sexual advance.[88] There's plenty in this that's contentious. Authors who propound this line of argument reject one of the framing assumptions in most free speech theory (including here, see §2), namely, that free speech principles are fundamentally about disallowing certain types of government action, and hence cannot underpin pro-censorship arguments.[89] They also favor a controversial view about how the social meanings and practices around sex operate, one that attributes a huge influence to pornography.[90] The part of this feminist critique that chimes with my analysis here, though, is simply its forthright—and by my lights correct and still underappreciated—insistence that free speech has to be about people's freedom to perform important speech acts; not merely to *say* things, but rather to *do* important communicative things (including, but not limited to, expressing opinions) with their words. A Subdividing Approach to free speech, which replaces a principle of free speech *per se* with principles protecting Freedom of Art, Freedom of Protest, and so on, is geared to accommodate this understanding of what communicative liberties are really for.

13. Conclusion: Two Kinds of Ideal Theory

Free speech theory has to circumscribe the domain of activities to which it applies. Subtractive Approaches aim to do that, by identifying the communicative acts that don't merit free speech's special protections, and extracting them out from the protected category of SPEECH. Different strategies for making these subtractions have been examined here, and in each case I've argued that Subtractive Approaches impede our attempts to define the protected class of communicative acts in a way that answers to the ARGUMENTS, that is to say, in a way that only

88. See Rae Langton, "Speech Acts and Unspeakable Acts" (1993) 22:4 Philosophy & Public Affairs 293; Rae Langton & Caroline West, "Scorekeeping in a Pornographic Language Game" (1999) 77:3 Australasian J Phil 303; Ishani Maitra, "Silencing Speech" (2009) 39:2 Can J Phil 309. The claims from MacKinnon appear at a number of points in her work, including *supra* note 63. Arguments to similar effect have also been made by critical race theorists defending the regulation of racist hate speech, e.g., Charles R Lawrence, "Crossburning and the Sound of Silence: Antisubordination Theory and the First Amendment" (1992) 37:4 Vill L Rev 787. However, the speech-act-theoretic framework that's integral in recent feminist philosophy on this issue hasn't been widely used in critical race theory in this area.

89. Under this analysis, free speech is understood as giving us positive reasons to restrict some kinds of communication, so as to enable or disinhibit other kinds of communication. Outside of the anti-pornography literature, this approach to free speech has been defended by mainstream theorists of constitutional law; see, e.g., Owen M Fiss, *The Irony of Free Speech* (Cambridge: Harvard University Press, 1996); as well as by proponents of classical liberal political theory; see David O Brink, "Millian Principles, Freedom of Expression, and Hate Speech" (2001) 7:2 Leg Theory 119.

90. For criticism of this analysis on both fronts see Daniel Jacobson, "Freedom of Speech Acts?" (1995) 24:1 Philosophy & Public Affairs 64; Leslie Green, "Pornographizing, Subordinating, and Silencing" in Robert C Post, ed, *Censorship and Silencing: Practices of Cultural Regulation* (Los Angeles: Getty Research Institute for the History of Art and the Humanities, 1998) 285.

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press

extends special protection to the acts that really merit it. What follows from this, however, is not just that we should switch over to an additive definitional method, but that we should reconceive the categories of communicative liberty that structure free speech policy frameworks. In place of the umbrella classification of SPEECH, we should recognize a plurality of narrower categories of protected communicative action, each of which corresponds with a particular family of free speech ARGUMENTS, and each of which is populated by only those communicative acts that are integrally involved in the social processes via which the values/ideals adverted to in the relevant ARGUMENTS are promoted or imperiled. The communicative acts that are elected to a privileged status, under this subdivided system of communicative liberty, should still be accorded robust protection against legal restriction, just as proponents of free speech orthodoxy recommend. But these protections, under the system that I'm recommending, aren't extended to anything that can be identified with a class of things as broad and nebulous as 'SPEECH'. Indeed, there's ultimately no method for defining SPEECH that fares well by the 'answering to the ARGUMENTS' criterion, because the very postulation of some reified, overarching class of communicative acts already misrepresents what the ARGUMENTS establish. We should use an Additive Approach to define categories of protected communication, then, but this methodological shift should be packaged with a—pluralized, ramified—reconfiguration of the categories we're trying to define.

But to whom exactly are all these recommendations addressed? The reader may question whether it's sufficiently realistic, on my part, to call for such a drastic renovation of the structure of free speech theory. I'll finish by speaking to this worry. *Ideal* political theory involves proffering an account of the conditions of justice to be aimed at, whereas *non-ideal* theory consists in discussing what we need to do in practice to promote those conditions, given present (unjust) realities.[91] Free speech theorizing is customarily conducted in the ideal mode. Theorists don't devise political tactics aimed at realizing a good and just free speech system; they try to say what such a system would involve in the first place. This is especially evident when it comes to the ethical underpinnings of free speech. Many values and ideals are threaded into free speech discourse, and while they aren't all necessarily at loggerheads with each other, nor is it possible for them all simultaneously to be the normative cornerstone of free speech theory. One way to idealize in free speech theory, then, is by articulating an ethical or political vision of free speech which cuts through this disarray, and nominates one particular value or ideal as *the* normative cornerstone of free speech, advocating adjustments to free speech policy to bring it into alignment with that

91. These definitions are borrowed from Simmons, whose definitions are based in turn on Rawls's (1971) original distinction between ideal and non-ideal theory; see A John Simmons, "Ideal and Nonideal Theory" (2010) 38:1 Philosophy & Public Affairs 5; John Rawls, *A Theory of Justice* (Cambridge: Belknap Press, 1971). There is another kind of idealizing methodological maneuver, which Rawls and some of his followers favor, which consists in positing principles of justice under the assumption that the agents governed by those principles will all fully accept and fully abide by them, i.e., the assumption of 'strict compliance'. When I speak of ideal theory here I'm not meaning to refer to this kind of 'strict compliance' method.

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press

vision. Such idealizing isn't merely tolerated in free speech theory, it's generally welcomed. As Seana Shiffrin says, articulating ethical foundations that are "independent of our actual historical tradition is essential to allowing us to assess which aspects of our historical tradition are worth valorizing and which should be amended", and it also "supplies the resources to engage in the … function of understanding our extant traditions in their best light".[92] Idealizing analyses of this kind are commonplace in the literature, regardless of the infeasibility of the recommendations they yield, and this is presumably because participants in the discourse think it's a worthwhile enterprise to try to envisage a better form of free speech politics, without direct regard to the implementability of the policies that would emerge from it. Insofar as that's the case, free speech theorists can hardly object, on grounds of infeasibility, to the kind of structurally idealizing free speech framework that I'm espousing here.

Of course it would be difficult at best for a First Amendment analysis to subdivide its protected categories of communicative liberty, and do away with the overarching classification of SPEECH, given the centrality of the term 'speech' within that framework.[93] It might be that the legal system which is best-placed to restructure its free speech policy in accordance with the Subdividing Approach is Australia's. Unlike other similar liberal democracies, Australia has no constitutional or federal bill of rights, although political movements aimed at implementing an Australian bill of rights have arisen repeatedly, and currently two states (the Australian Capital Territory and Victoria) have state human rights acts.[94] If in future Australia were to implement a constitutional bill of rights with a system of judicial review, there is no in-principle reason why this bill could not structure special communicative liberties in Australia by enshrining

92. Seana Valentine Shiffrin, "Methodology in Free Speech Theory" (2011) 97:3 Va L Rev 549 at 550. Instead of positing ethical bases *independent* of First Amendment doctrine, to try to improve or clarify that doctrine, it's at least *as* common an approach, among American legal theorists, to seek to discern the core values operative *within* First Amendment doctrine. On this issue I favor the position that has its roots in Meiklejohn, *supra* note 15, on which the ideal of participatory democracy—roughly: the opportunity to participate in collective self-government—functions as the First Amendment's normative core; I favor this position because I agree with Weinstein that, if any ideal can be ascribed a central position in this jurisprudential tradition, participatory democracy "is the only contender that the case law does not massively contradict"; James Weinstein, "Participatory Democracy as the Basis of American Free Speech Doctrine: a Reply" (2011) 97:3 Va L Rev 633 at 643. But it is consistent with this to think that a broader set of ethical ideals *ought* to ground free speech policy, and that the First Amendment isn't structurally conducive to implementing policies that answer to a broader set of ideals.

93. Schauer has spoken of a "multi-valued theory" which treats "the free speech and free press clauses of the First Amendment as the umbrella under which are located a number of more or less distinct separate principles, each with its own justification, each directed towards a separate group of problems"; we might, he says "have several First Amendments"; Schauer, *supra* note 4 at 1303. Horwitz's discussion of a multifaceted, institution-oriented approach to the First Amendment examines how something like this might be achievable within the institutional constraints of First American doctrine; see Paul Horwitz, *First Amendment Institutions* (Cambridge: Harvard University Press, 2013). Although there would be plenty of complications and difficulties in practice, subdividing free speech within a First Amendment framework may still be possible.

94. See Jeffrey Goldsworthy, "Introduction" in Tom Campbell, Jeffrey Goldsworthy, & Adrienne Stone, eds, *Protecting Rights Without a Bill of Rights: Institutional Performance and Reform in Australia* (Aldershot: Ashgate, 2006) 1.

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press

multiple principles of communicative liberty—Freedom of Art, Freedom of Protest, Freedom of Inquiry, etc.—instead of one single overarching principle of Free Speech. It's also worth noting, in this connection, that the Conservative government elected in the United Kingdom in 2015 has said it plans to replace the U.K.'s 1998 Human Rights Act (which brings the rights articulated in the European Convention on Human Rights into U.K. law) with a new British Bill of Rights.[95] Free speech rights in the U.K. currently receive some measure of protection under the common law, and while a new British bill of rights wouldn't have anything like the same degree of influence as the First Amendment in U.S. constitutional law (because of the British courts' much more limited ability to alter parliamentary legislation), nevertheless a new British bill of rights could articulate and endorse a subdivided conception of free speech, and thereby have a major effect on how free speech policy evolves in the U.K., via both parliamentary and judicial action. This is all speculative, obviously. Whether and how a Subdividing Approach to free speech theory could be put into action in any real-world jurisdiction any time soon, is a question whose complexities cannot be adequately addressed in this paper. The point that I was making above, to reiterate, is that free speech theorizing routinely sets itself outside the legal institutions that it is speaking to, and proposes reforms that couldn't be realized without an overhaul of the institutions in question. Restructuring free speech theory in order to subdivide the category of SPEECH would be a major overhaul indeed, but that's not a good reason to dismiss the case for this proposal.

95. See *The Conservative Party Manifesto 2015* at 60, online: https://www.conservatives.com/manifesto.

https://doi.org/10.1017/cjlj.2016.20 Published online by Cambridge University Press