

発話の3つの害の為し方と非規制的アプローチの諸相 オンライン・ヘイトスピーチ規制の内在的制約

東京大学大学院情報学環
准教授 永石 尚也
n.nagaishi@iii.u-tokyo.ac.jp

キーワード: ヘイトスピーチ、尊厳、言語行為論、モデレーション、政治の状況、アイデンティティ政治

※ 別稿「発話の3つの害の為し方と非規制的アプローチの展開 オンライン上のヘイトスピーチ対応と法的介入のグラデーション」(本多康作ほか編『差別・ヘイトスピーチ・法』(近刊)所収)と併せて参照されたい。以下ではこれを「論考[展開]」と略記する。

はじめに 「ヘイトスピーチは罰すべきか Is Hate Speech Punishable?」を問う前提を分節化する

共約不可能な断絶が介在しているため、世界への欲望、世界の悪、世界を求める存在が通過、架橋、地峡、交流、翻訳、転義、転移の試みを提示し、強要し固定化しようとするのですが、それらの試みはことごとく潰えます。世界は存在せず、ただ島々だけが存在するのです。私はツェランの短い見事な詩の最後の一節「世界は消え去っている、私は君を担わなければならない」を数多くの方向性で解釈しようと努めるのですが、その一つがこれです。

ジャック・デリダ『獣と主権者II』第一回[2002年12月11日] (西山・亀井・荒金・佐藤訳)

世界が存在しないような場所で、世界がここにもあそこにも存在せず、遠くに、向こうに無限に離れているということ。君と共に、君を支えながら私がなさなければならないことは、まさに一つの世界がある、ただ一つの世界がある、さもなければ一つの公正な世界があるようにすることだ、ということ。あるいは、あたかもたった一つの世界しか存在しないかのようにし、世界を世界へと到来させるようにし、君に対して、君にそれを与えるために、君に向かって、君の方へとそれを届けるために、君にそれを送付するために、私が世界を世界へと到来させるかのようにすること——あたかも現在そこに存在しない場所に世界が存在するかのように。

ジャック・デリダ『獣と主権者II』第十回[2003年3月26日] (西山・亀井・荒金・佐藤訳)

本稿では、**オンライン上でなされるヘイトスピーチの性質とその害の特性を、夙に近時において新たな発展を遂げてきた言語哲学上の発語媒介行為・発語内行為の観点¹から明示化**

¹ ヘイトスピーチという言語現象に対してその言語内行為性から分析した代表的な論考には、Rae Langton “Speech Acts and Unspeakable Acts,”(1993)やそれを引き継ぐ Maitra and McGowan “Subordinating Speech”(2012)がある。これらを筆頭に、現在に至るまで主としてフェミニズム哲学の背景の下で各種の分析が進められてきた。

とりわけ近時、Mary Kate McGowan による文字通りに両義的な名前を持つ著作 *Just Words: On Speech and Hidden Harm* (2019) を通じて広範な領域から論じられるようになってきている。直近では、特に同書をめぐっての論文集 *Australasian Philosophical Review*, 5(2) (2022) の Sarah Sorial & Sam Shpall, Mary Kate McGowan, Ishani Maitra,

した上で、その許容性と望ましい（刑事）法的介入の態様を判断するにあたって考慮すべき事由を明らかにすることを目指す。しかし、前者の手法的な部分（発語媒介行為・発語内行為による分析）については別稿論考[展開]に譲り、むしろそうした議論によって「(法)規制」による介入を正当化しようとする営為が見落としている議論の前提を明らかにするべく、本稿では専ら後者に力点を置いて分析を施す²。後者の主軸となる議論は、いわゆる表現の脆弱性に発する議論を拡張したものである。すなわち、規制に晒された際のヘイトスピーチ（類似の発話）の脆弱性に加え、ヘイトスピーチ（類似の発話）に晒された際の対抗言論の脆弱性の双方に対して、発話上の沈黙化や抑圧等の「悪しき」効果がいかに生じるのかにかかるとして既存の議論を、オンライン空間の諸特性に照らして分節化し、適用範囲を画することを目指す。いわゆるコンテンツモデレーションなど、発話そのものへの不可視化や沈黙化が、国家という規制主体以外のデジタルプラットフォームによる準-規制作用を介して可能となるオンライン空間においては、従来のリアル空間での発話内効果・発話媒介効果による脆弱性の議論の射程は自ずから限定されるとともに、オンライン空間の特性に応じた新たな非規制的で協働的なアプローチを開拓する必要性が浮き彫りになる。

他方で、一般的な表現としての価値性の高低に発する議論、すなわち民主主義ないしは思想の自由市場を支える価値の毀損等なんらかの（発話にとって）外在的な目的から法的な保護が及ぶべき表現を拡大的に解釈（あるいはその低価値性ゆえに制限）する議論が対象にしてきたものと、本稿の検討の手順が異なるものとなることは、あらかじめ断りを入れておかねばなるまい。簡潔にこの見解を敷衍するならば、以下の通りである。

一般にある表現が一定の害の発生に関連する場合であっても、当該表現が直接に特定人への害をなす場合を除いては、多様な意見への寛容の観点から許容されるのみならず、特定

Robert Mark Simpson, Jeremy Waldron, Caroline West らによる各論稿及びそれらへのリプライは、論考[展開]においても第2節を中心として明示的に参照することとなる。

² 本稿は2022年3月の科研費報告会「オンライン（特にSMP）上のヘイトスピーチ対策と法的介入のグラデーション」（司会：本多康作）を修正の上で論文化したものであり、法領域の中でも（刑）罰の執行可能性に限定して論じることとなる。したがって、例えば公の施設の利用可否をめぐる諸課題や、SNS型のデジタルプラットフォームによるコンテンツモデレーションやアカウント削除等をめぐる諸課題等、多くの論点は別稿に譲らざるを得ない。この限定はもちろん国家による権限行使における刑罰という制裁の特殊性を前提としているが、特に同報告直前の2022年2月に最高裁が示した事例とも関連する。すなわち、「大阪市ヘイトスピーチへの対処に関する条例」（2016年）に基づく公金支出（審査会委員報酬支出）の合憲性が問われた住民訴訟、最判令和4年2月15日において、その合憲性の理由の主たる部分として、表現を制約する手段としては拡散防止措置としての掲示物等の撤去・オンラインコンテンツの削除要請及び認識等公表など、刑罰に至らない手段ゆえに必要最小限度の制約であるとして憲法21条1項に適合し、許容されるところと関連する。更なる比較対象として、本邦でも唯一例外的に罰則（12条）を規定する「川崎市差別のないまち人権条例」において、罰則対象をあくまでもリアル空間における程度の甚だしい行為へと絞り込むとともに、手続上も違反者への勧告、勧告に従わない場合の命令、命令に違反した場合の公表と罰則（罰金）という一連のプロセスを整備している点も参考になろう。また、読者による検索の便宜のために補足するならば、前掲注1)の論文集にも寄稿しているRobert Mark Simpsonを、これに先だつこと2年前の2020年2月に本邦に招き、まさに「Is Hate Speech Punishable?」と題して開催した科研費研究会（2020, 司会：永石尚也）に由来するものでもある点も備忘的に記す。

なおこの時のR. M. Simpsonによる報告「Is Hate Speech Punishable?」は論文として、近刊『差別・ヘイトスピーチ・法』（2023 発刊予定）第二部に邦訳「ヘイトスピーチは罰すべきか、罰に値するか」（永石尚也・本多康作 訳）として収められる。同じく所収されている他の参加者報告を論文化したものと併せ、参照されたい。

の場面における専門性・有用性等から（条件付きで）必要とされる場合ももちろんある。仮に、有用性が乏しいと広く目され、あるいはその不品位性が広く共有されている場合であったとしても、古くは後述する J. S. ミルが述べたように、現時点における我々の認識を疑いうるものとするのは社会にとって有益である。さらに、より重要なのは、成熟の達成を期し難い我々市民一般の多様性と認識的な能力欠如とに鑑みれば必ずしも事前に「適切な発話」を定型的に要求することは事実として困難であるとともに、規範的に見ても市民の中でも特に負担を強いられてきた属性の者にこそ「不品位」とマジョリティから目される発話を積極的に促すことは有益であり、許容すべきである。反対に、こうした表現への規制は、むしろ常に不合意に満ちた「政治の状況」を不可視化し、新たな政治的争点形成を阻害する懸念さえもある。おおよそこのように、主張されることとなるだろう。

これらの主張の妥当性については、本稿の末尾でも戻ることとするが、しかし本稿の主たる関心は、そうした表現の特性に着目した価値評価に先立ち、まずはオンライン上のヘイトスピーチの持つ害の形式的な特質につき、当該害の発生・拡大を社会状況・技術的環境に相関する形で明らかにすることにある。すなわち、ヘイトスピーチの持つ害が対抗不能な形で現実化する条件を当該発話の言語行為上の特性から明示化する前提として、オンライン空間の諸特性を明らかにすることが本稿の対象である。

以上を踏まえ、本稿の議論は下記の順に進む。

第 1 節では、オンラインヘイトスピーチが集団への侮蔑的・差別的発話を介してもたらす害とされるものについて、集団的属性と個人に対する害との関係及び非対面状況におけるオンラインヘイトスピーチの言語的特性を介した害の構成・発生について整理を施す。これにより、リアル空間におけるヘイトスピーチとの議論上の差異を生み出す論点を取り出す。

第 2 節では、上記の害の特質を踏まえた上で、オンラインヘイトスピーチの可罰性を問う前提を明らかにするべく、害の発生機序と（刑）罰の執行との間の距離から、可罰性概念と（刑）罰の執行を基礎付け、限定する法内在的な制約の前提を整理する。第 3 節では、より具体的に法内在的な制約としての害と規制とのリスクトレードオフを法的介入・不介入の場面ごとに分類する。

その上で、第 4 節では、害に着目する本稿の歩みとは異なる言論の自由市場論に発する議論蓄積の中に本稿で見てきた発話内的・発話媒介的アプローチを接続する可能性とその位置付けを示し、論考[展開]への接続を期することとする³。

³ なお、専らオンラインヘイトスピーチの害を手法面で分析することを目指す論考[展開]での議論を本稿と関連する限りで記せば、以下の通りである。その冒頭では、具体的にヘイトスピーチによる害を因果的害と構成的害とに区分した上で、その発生機序を発話媒介的アプローチと発話内的アプローチのそれぞれから分析することを示す。まず第 1 節では、ヘイトスピーチがもつ発話媒介効果に着目した分析を、オンライン/非オンライン状況の比較において施す。この結果として、とりわけオンライン上においては、発話文脈の多様性と規制主体の多層化、そして発話者と宛先となる対象者との時間的・空間的・関係的「距離」により、規制の基礎となる明確な害悪は疑わしいことを導く。

次いで第 2 節では、発話内行為としてのヘイトスピーチが持つ効果に着目した McGowan による分析を、そのアイデアの元となった David Lewis の scorekeeping 論とその現代的解釈を与えた Derek Ball の議論によって修正を施しつつ、オンライン状況へと適用する。ヘイトスピーチがオンライン上においてなされる場合、時間的・空間的・関係的な

なお、本稿全体で主として参照される諸理論の多くは、いずれにおいてもその構成要素の一部を、明示的に J. L. Austin の記念碑的な講演「いかにして言葉でものごとを行うか」(1955) に負っている。言語行為論そのものは、その後、言語哲学、フェミニズム哲学、フランス現代思想、社会学、歴史学等の人文・社会科学諸分野の壁を横断して拡散してきたとも言えようが、ヘイトスピーチという主題をめぐって各々の理論的資源を持ち寄り協働している様を、本稿の裏側には透かし見ることもできるかもしれない⁴。

「距離」により、会話において許された状況的規範の定立・適用は、多くの場合十全に作動しない。むしろオンライン上においては、規範定立の不発を前提に、当該発言への非難や通報など対抗的な議論喚起もまた惹起することで、当該社会における差別的視線と共に「差別への(対抗的)視線」をも可視化することになることが示唆される。

以上の発語媒介的アプローチと発語内的アプローチに基づく分析を踏まえるならば、文脈に応じ、多種多様ななされるヘイトスピーチの持つ害の不確かさが直接的な法規制の根拠としての適格を欠く要因となる。しかし、オンラインヘイトスピーチに属する発言がなされた場面においてそれへの対抗言論を行使する際、(解釈上その他) 諸資源の不均衡に鑑みた格差の害という3つの害をも持つだろう。論考[展開]第3節では、共生に向けられた環境形成の指導理念としての社会的な「尊厳」の概念を改めて分析することで、法の役割を改めて位置付ける。具体的には、平等な地位と自尊心、そしてシティズンシップという社会的尊厳を構成する三層からオンライン上におけるヘイトスピーチへの対抗言論アプローチを位置付けるアプローチによって、共生に向けられた議論環境を支える法の役割を示すこととなる。

⁴ 以下、読み飛ばしうることを前置きした上で、エピグラフに掲げた講義の日付周辺を辿ることで、本稿全体の解題に代えよう。同引用部は、ジャック・デリダによるダニエル・デフォー著『ロビンソン・クルーソー』及びマルティン・ハイデガー『形而上学の根本諸概念』における Walten 概念への読解に、その大部が費やされた最終講義「獣と主権者 II」(2022-2023) に由来する。もちろん政治的に読めば、これが執筆当時におけるアメリカの過剰主権と孤立主義とを狙うものであることは確かであろう。しかしデリダによる読解対象はそれに尽きるものではない。まさしく2つの引用部とともに掲げられているパウル・ツェランの詩“GROSSE, GLUHENDE WOLBUNG”の末尾の一節、「世界は消え去っている、私は君を担わなければならない」が主軸であるとさえいえる。この引用箇所は2022年12月から2023年3月の時間をもつが、そこ挟まれた2003年2月5日のハンス＝ゲオルグ・ガダマー1周忌追悼講演「雄羊」で、ガダマーによるツェラン読解に重ねられる形で最も詳細に検討された。2002年3月のガダマーの死後、デリダは3周忌の日付を待たずにその生涯を終えることとなる。その僅かな二つの死の間隙を縫ってなされた上記の読解は、まさにある(亡)者から僅かに生き延びた証言者として、言葉を持たぬ(亡)者に代わって、しかしその(亡)者に成り代わることなく語る術を、その一節の解釈として提示していた。

この提示の仕方は全く偶然ではなく、デリダによる“GROSSE, GLUHENDE WOLBUNG”の読みそのものであった。(喪を介した) 解釈による他者の「取り込み、記憶の内化、理想化」を可能な限り排した上で、かつての言葉を相続し、新たな言葉をまだ持つ者の義務を、デリダは取り出そうとしていたからである。引用部では、この「取り込み、記憶の内化、理想化」を(喪のみならず)力による変更に対しても同様に、すなわち正統化された実力(主権者)に由来するものであれ剥き出しの暴力(獣)に由来するものであれ、あらゆる「取り込み」一般へと抵抗するものとしてパラフレーズしていたものとして捉えられよう。我々は獣でも主権者でもあり、不可避的に暴力を行使する(常に赦しを乞う)側に立つと同時に、時に応じてそれらに晒された羊の側で抵抗を試みるような両義的な存在でしかありえない。

死によって今はもう言葉を発せない者、あるいは暴力や監視によって(例えば口に出した隠語(シボレート)の響きによって出自を明らかにしてしまうがゆえに)口をつぐんでしまう者のために、デリダは翻訳不能な「秘密」の言葉の領域を開拓していた。そうした言葉を共有しえない他者と「共住」するための世界は、上記の別離の只中の誓約(義務)を通じた「かのように」によってこそ辛うじて現出する(西山雄二(2014)「世界の終わりの後」思想(1088), 118頁)。少なくともそこにデリダは「信」をかけていたことは確かであろう。

本稿では上記の引用部を導きの糸とし、オンライン上のヘイトスピーチに向き合うために、一見すると逆説的にも思えるかもしれない「赦し」の空間というモチーフへの接岸を期したい。辿り着いた島の内、暴力が暴力としてすら力を失い自由が自由の意味を失う孤独の中で、ロビンソン・クルーソーは既定の世間とは隔絶された他者の声を聞く。無主地と考えた傲慢から発されたロビンソンその人の声に暴力が根ざしているのか、ロビンソンに対峙した他者の声が暴力の片鱗を宿しているのか、それを決定しても仕方はない。どちらにも暴力の兆しはある。そして、その声の主権者によるものであれ獣によるものであれ、危機を煽る憎悪の言葉に立ち向かうのは、ある特定の観点から正統化された実力であってはならないし、別の剥き出しの暴力であるところの、憎悪に対抗するカウンターとしての憎悪の言葉であってもならないだろう。むしろ、ターゲットにされた「羊」と共に在る「かのように」生きる義務を引き受けた我々の「共住」の言葉に他ならない。そしてさらに問題となるのは、ツェランの“GROSSE, GLUHENDE WOLBUNG”の末尾の直前には「何に向かって彼は突進しないのか」と置かれていた通り、ここで対象を名指すことなく反旗を翻す「羊」が、

現実の物理空間で位置を占める「誰」のことを指すのが開かれた問いであることに他ならない。かのような信を置きつつ、以下、本論へと入ろう。

1、ヘイトスピーチが持つ害の種別と、規制による害の種別

1.1 リアル空間におけるヘイトスピーチの性質とそれを取り巻く様々な害

(1) リアル空間におけるヘイトスピーチのもたらす害

改めて確認するまでもなく、そもそも特定の民族的・宗教的属性を理由とするヘイトクライムや差別的取扱いが、特別にそれらを括り出す規定を持つか否かにかかわらず、広く法的に禁じられることは疑うべくもない⁵。そうした物理的暴力や取扱いに至らないヘイトスピーチであっても、特定人に対する侮蔑・中傷や威圧、脅迫、プライバシーや名誉を脅かす虚偽の事実の流布、暴力の扇動等の行為類型に属する場合には、同様の禁止が働く。

しかし、こうした行為群の近傍に属する発話群のうち、何をヘイトスピーチに属するものとし、上記行為群と同様に禁じられるべきものとするかについては、理論的にも実践的にも、その害を弁別し、慎重な検討を要するべき多くの理由がある。

ここで冒頭に挙げた発語媒介行為・発語内行為による害を分けて論じることが一つの助けになる。詳細は別稿論考[展開]に譲るが、発語媒介行為アプローチとは、端的に言えば、「それをいうことによって (by) ものごとをなす」その発話効果から、行為としての発話の特徴を把握するアプローチを指す。例えば、偽情報拡散類型は、基本的にはそれが惹起する信用毀損・業務妨害（刑法第三十五章 信用及び業務に対する罪等）や、世論形成⁶・選挙等を含む広義の政治過程に対する効果（公職選挙法 第十六章 罰則等）⁷、公正な競争（独占禁止法、不正競争防止法）や健全な金融市場の毀損（金融商品取引法第六章 有価証券の取引等に関する規制等）等、表現外在的な社会的害に依存する形で「害」を実現するのであり、単に内容が真実と異なる情報であるとの表現内在的要素だけで「害」をなすものではない⁸。

⁵ 京都地裁令和4年8月30日（非現住建造物放火罪）は、行為及び結果としては通常の犯罪と同一であったとしても差別的な動機が動機の発現が量刑上も考慮された事例として紹介される。判例解説として、金尚均（2022）「一連の犯罪について差別的動機が量刑上考慮された事例」を参照せよ。

⁶ 辻田真佐憲・西田亮介（2021）『新プロパガンダ論』（株式会社ゲンロン）。なお両氏も想定しているだろうように、「耳を傾けすぎる政府」問題は決して日本特有の事態ではなく、SNSによって政府への賛否を含む意見が可視化される場所では見られる現象である。例えば、権威主義的立憲主義と目されるシンガポールにおいても同様の事情は観察される。この点につき、永石尚也（2021）「公共空間における情報識別と警察活動についての覚え書き」第一節を参照せよ。

⁷ 曾我部真裕（2020）「インターネット上の情報流通と民主政」法学セミナーNo.788などを参照せよ。本邦では広告規制などを介した不均衡の是正は議論されてきたが、表立って偽情報対策が現行法上の規律には乗っていない。これは当然ながら、政治過程そのものが一見非科学的に見えるものであれ特定の利益や信条の発露を真っ向から否定するものではなく、それを前提しつつ正当性・正統性の調達の回路を複層的に構築する営みであるためである。無論、量的な限度や誤導的な情報の害悪がある以上は、その介入の限界についての問題提起は種々になされている。

⁸ ハーマン・カペレン&ジョシュ・ディーバー（2022）『バッド・ランゲージ：悪い言葉の哲学入門』（勁草書房・葛谷潤、杉本 英太、仲宗根 勝仁、中根 杏樹、藤川 直也 訳）第3節3-2-2、Faulkner(2007)から引かれている箇所での例を見ることができる。ナチスの追手からある者を匿っている家主が、誰かを匿っていないか、と尋ねられた時に「いいえ」と嘘を述べることは、結果としてはその者に利する蓋然性が高いだろう。さらに真実を述べつつ嘘をつきうることを示す例としては、上記の例で追手に家主が尋ねられる直前に匿っていた者が窓から逃げ出していた、という例を提示できる。

さらに、自らが真ではないと信じる事実を他者に真だと信じさせる点で、嘘と欺きは共通点を持つが、いずれもその

逆に真実を述べることで「害」をなす場合がありうるのは、名誉やプライバシーに関わる真実を伝えることを念頭におくだけでも明らかであるだろう。扇動・教唆類型でも、基本的にはそれが惹起する非言語的な暴力や暴動等の危険が実現化するか、差し迫ってなされる蓋然性が高い状況においてその「害」を実現する。相手にその決意を生じさせただけで「害」を作り出すというのは破防法等のごく例外に属する場合には存在しなくはないものの、理論的に大いに異論含みであり、かつ、その罪としての成立は当然に限定的である。ましてや単にそれが他人を煽る文面でなされることのみで「害」をなすとするのは、およそ賛同を得がたい見解であろう。

以上のように、言葉・情報によって害をなす既存の類型に属する発話は、その言葉によって因果的に接続された「害」をなす類型であり、その拡張を(特別法によって)行うにせよ、発話とは独立の外在的な価値(害)によって補強されているものといえる。これに対して、脅迫による恐喝・強要、侮蔑・名誉毀損などは、まさにある状況下において特定の言動を表示することそれ自体で、いわば「それをいうことにおいてものごとをなす」ために発話自体で「害」を発生させる行為類型である。これは発語媒介アプローチと区別される形で、発話内アプローチとして、論考[展開]の3-1 発話内行為の箇所において詳述する。

(2) 規制のもたらす害

さて、以上を確認した上での更なる問題として、ヘイトスピーチとその他の類型を峻別する基準の問題がある。(直接に該当しないながらも)上記行為群行為群と共にヘイトスピーチがなされることが傾向としてあるとしても、また当該表現が職場など離脱困難な場所的限定や地位関係・力関係においてなされる場合にはそれらが近接する場合がある⁹としても、ヘイトスピーチに依存しない行為の害とヘイトスピーチそのものによる害とをまずは分離する必要がある点が挙げられよう¹⁰。仮に、それらの害が複合的に生じる蓋然性が高いと考

後の事態の推移と相手からの信頼への違背継続などに対する規範的評価という事後的事実によって、その「害」が把握されることになる。

最後に、フィクション作品において内容上妥当する「フィクションにおける真実」(デイヴィッド・ルイス)は、その真偽判断において参照されるべき世界が現実世界と異なる特定の虚構世界にあり、その信念を抱くことは単なる現実世界との不一致という意味での偽の信念を抱くこととは異なる点もここで指摘できる。こうした「フィクションにおける真実」の多層性は、自身とは異なる他の人生や final vocabulary に対する尊重や寛容を育む点においても有益なものでありうるものと思われる(リチャード・ローティ(2000)『偶然性・アイロニー・連帯』齋藤純一・山岡龍一・大川正彦訳(原著1989年)153頁)。この点は、ウォルドロン⁹の議論に即して、論考[展開]でも詳述する。

⁹ 大阪高裁令和3年11月18日はリアル空間の職場でなされた文書配布について賠償請求が認められた事例であるが、「●●人は死ぬ」といった文言やその思考を「野生動物」に喩える表現をヘイトスピーチと認定した一方で、「売国奴」などの表現はヘイトスピーチには当たらないが差別を扇動し、分断を強化する侮蔑的発話であるとしている。その上で、こうした文書の配布あるいはその放置がともに職場環境配慮義務に違反するものとし、その同時発生性を認めている。判例解説として、武井寛(2022)「使用者による差別的言辞文書配布(ヘイトスピーチ)と職場環境配慮義務/ (大阪高裁令和3年11月18日)」を参照せよ。なお本事案において「労働条件」に職場環境配慮を含まなかった点を問題視するものとして橋本陽子(2022)「使用者による国籍差別的言動(ヘイトスピーチ)の違法性——フジ住宅事件」がある。

¹⁰ Matthew Kramer (2021), Freedom of Expression as Self-Restraint 第2章 [Components of a Theory of Freedom of Expression]では、統治上課される自制として、4つの中立性、すなわちコミュニケーションの形式、内容(主題・観

え、規制根拠とする場合には、まさにその分かれ難さを生ずる場面を厳格に特定する必要があるためである。実際、環境に即した対応が情報類型ごとに、継続的な制度的整備とその見直しの中にある¹¹のはこれを示しており、事業者の自主的取り組みを促すとともにその過剰さによる不利益を回避させる統制された介入の下¹²、人権及びデュープロセスや理解可能性・説明可能性等の諸原則¹³に基づいてソフトな秩序形成を促進する「協調」が模索されている現状にある¹⁴のも同様である。

この根拠は何よりも、規制可否とその条件・態様を論じる文脈における「ヘイトスピーチ」についての法の制定・解釈・適用それ自体が、他ならぬ国家による権限行使を前提とした営みであることに求められる。つまり、上記の中心事例が生じさせる害や行為類型が折り重なって生じる傾向性等を前提としつつも、当該中心事例からの場当たりの類推や傾向性による拡大処罰に至らぬよう、適用段階における手法の統制はもとより、それに先立つ解釈を恣意化させることをも避ける必要があるのである¹⁵。

点)、主体によって取扱いを変えないという中立性と、行為との結びつきが偶然的である(行為からの独立性をもつ)コミュニケーションに対して干渉を控えるという中立性が要求されるとしている。

¹¹ 総務省「プラットフォームサービスに関する研究会 第二次とりまとめ」(2022年8月)における違法・有害情報対策とともに、利用者情報の取り扱いもあわせて参照せよ。また、発信者情報開示に関して、総務省「発信者情報開示の在り方に関する研究会最終とりまとめ」(2020年12月)を参照せよ。例えば前者では、権利侵害情報についてはプロバイダ責任制限法及び関係ガイドラインにより、その他違法な情報についてはインターネット上の違法な情報への対応に対するガイドラインにより、違法ではないが有害な情報については違法・有害情報への対応等に関する契約約款モデル条項により、青少年に有害な情報については青少年インターネット環境整備法によるフィルタリング等によって対応することが整理されている。

¹² やや異なる例ではあるものの、横大道聡『現代国家における表現の自由一言論市場への国家の積極的関与とその憲法的統制』(2013年)第8章を参照せよ。そこで論じられているように、フレデリック・シャウアーは論文“Principles, Institutions, and the First Amendment”(1998年)において、芸術、図書館、大学、公共放送ネットワーク等の様々な文化制度について、制度それぞれの規範・法理を検討する必要性を論じ、制度特定の修正1条論の構築を提唱していた。こうした文化制度が本来的に果たす機能の侵害を(部分社会の再来の危険を持つ「専門家の自律」保障ではなく)「公共討論への違憲な介入」という枠組みから捉えることを、横大道は主張する。なお、そこでは「あるべき公共討論」が指図されるのではなく、制度の本来的役割への阻害が「あるべきではない公共討論」への転化として捉えられている点にも、今後の検討における本稿との接続を期待することができる。

¹³ 「サンタクララ原則—コンテンツモデレーションにおける透明性と説明責任について」(2018策定、2021改訂)でも、国家が事業者に代わってモデレーション判断に関与することは排除される一方で、事業者による恣意的な判断を回避すべくアカウントビリティ確保のための原則が列挙されている。あわせてアメリカにおける通信品位法230条(c)(2)(A)の免責のための要件をめぐるオンラインの検閲の防止に係る大統領令(2020年5月28日)及び司法省勧告(2020年6月17日)を参照せよ。同免責の射程を限定する各種否認法理については、平野晋「免責否認の法理(『通信品位法』230条)」総務省・情報通信政策研究所(2014年)を参照せよ。

¹⁴ 本邦において特徴的な例としては、川崎市における「川崎市差別のない人権尊重のまちづくり条例」における、インターネット上のコンテンツに対するインターネット拡散防止措置(プロバイダ等への削除要請)がある。法令上の効力として削除要請に応じない場合の制裁は伴わない点では、その効果としてはヘイトスピーチのない社会の実現に向けた市の認識を提示する象徴的效果や表現内容等の情報提供機能に留まり、また手続的に学識経験者により構成される審査会の意見を踏まえるものとして評価できる。他方でこうした「要請」にこそ、潜在的な個々の表現者にとっての萎縮効果を排した明確性とアカウントビリティが求められる点は、本文記載のとおりである。法における正統性確保を支える「誠首」とは異なる正統化原理の動揺については、興津征雄「正統性の構造分析—行政国家の正統性を手がかりに(上)」法律時報93巻1号及び同論考「(下)」法律時報93巻2号を参照せよ。

¹⁵ ジャック・デリダ『獣と主権者II』(白水社、西山雄二・亀井大輔・荒金直人・佐藤嘉幸 訳)。冒頭のエピソードに引いた一節に掛けて言えば、次のようになる。剥き出しの(獣の)暴力を排除する限りにおいて主権者の権力は統制されつつ呼び出されてきたわけだが、しかし、技術的・環境的变化によりある行為の可能性及びそれがもたらす利益・危害の可能性が拡張した際、それがもたらす害のみを取り除くことは困難に直面する(デュアルユースと同様である)。だか

これはオンライン空間においてヘイトスピーチがなされる場合にも当てはまる。一般に、ある表現類型 S のもたらす害 H について新たな法規制 R を求めることは、当該規制作用への制約を制度的均衡を目指して実行する装置なくしては、制定法の解釈・適用を通じて実現する権力に対して概括的に正統化を付与する（いわば正統性のバックドアを開くこと）と同義である。ましてやオンライン空間の場合、発話そのものに対する介入が技術的に可能である。そのため、 S にあたる個別表現 s の選択段階で恣意性がある、あるいは少なくともその不信が生じるならば、表現類型 S を超えて個別表現 s の生産・流通は阻まれようし（萎縮効果）、害 H の状況に依存した具体化であるところの個別の害 h の選択段階で恣意性がある、あるいはその不信が拭えなければ、リアル空間以上に萎縮を働かせる事態を招くだろう。

いわゆる表現の脆弱性に即して、コミュニケーション中立性及び主題・観点・発話者の中立性が規制主体に要求される¹⁶のは上記事情のためであったが、オンライン空間におけるヘイトスピーチを規制しようとするれば、（オンラインに流通する情報それ自体としては物理的な威力・暴力行使には至らない以上）基本的にはまさに文字・画像・映像等の情報内容に即して規制可否の判定がなされるより他はなく、上記の中立性要求に否応なく抵触する。この点に鑑みれば、抑止の必要性を満足させるための介入を限定する内容・手続上の規律と、それらへの（司法）審査がともに厳格になされるべきことへと帰結しよう¹⁷。

規制をめぐる議論におけるこの事情のためにこそ、他ならぬヘイトスピーチに固有の問題を他の差別的・差別助長的な行為類型と（鶴的）に混ぜ込むことなく、明確に区別して取り出す必要がある。殊にオンライン上のヘイトスピーチの害を論じるためには、次節以降で見ていくようにオンライン空間のコミュニケーション上及び空間上の特性として差異を踏まえて、オンライン上のヘイトスピーチに属するべき発話類型を同定する必要があるのである。

1-2. 規制をめぐる困難① 定義上の文脈鋭敏性

（1）ヘイトスピーチ規制の多様性と不一致

からこそ、害が現実化する手前において、内なる暴力の「徴候」を介入可能なトリガーとするならば、主権者の panvasive な権力の供養へと自らを捧げることとなる。自己保全の力が自己破壊へと転化する民主主義の自己免疫的事態に至るのは、このためである。

¹⁶ 前掲注)10 Matthew Kramer (2021), Freedom of Expression as Self-Restraint 第2章 [Components of a Theory of Freedom of Expression]

¹⁷ 最判令和4年2月15日の「大阪市ヘイトスピーチへの対処に関する条例」(2016年)に基づく公金支出(審査会委員報酬支出)の合憲性が問われた住民訴訟の事例では、人種差別的な意識の助長をもたらす言動についての抑止の必要性から、その詳細に踏み入ることなく規定の合憲性を導いている。しかしこれが法的強制力を伴う刑罰を付加するケースにおいて妥当するかは疑わしく、にもかかわらず法的介入を要求するヘイトスピーチについての限定を明示的には施さなかった点については不足がある。この点に関連して、本邦における裁判所の審査手法をヘイトスピーチ規制に対して詳細に検討したものとして、Keigo Obayashi (2021)“The free speech jurisprudence in Japan: the influence of comparative constitutional law,” in (ed.) Shinji Higaki, Yuji Nasu, Hate speech in Japan : the possibility of a non-regulatory approach に詳しい。

さて、オンライン上のヘイトスピーチに属する発話類型を同定し、それが生じさせる多種多様な害とその発話類型を明示的に区分して論じるためにも、先んじて、比較対象としてのリアル空間を前提とした既存のヘイトスピーチをめぐる議論蓄積と現実の規制（とその拡散）状況を把握することは有益であろう。

そもそもヘイトスピーチとして把握されてきた中心的な事例、例えば比類なき道徳的悪としか断じ得ないジェノサイド肯定の発話やホロコースト否定論を基礎とする一群の民族的・宗教的集団への蔑視的・排除的発話については、歴史的経緯及び社会状況に照らし、その発話自体がもたらす害が複合的・重畳的に折り重なり、社会における許容性を結果として超えるとのコンセンサスが築かれてきた。すぐ下で見るように、(刑)罰の執行を含めた法規制その他の対応も各国においてこれまでも実践として積み重ねられてきたし、民族的・宗教的集団に対する憎悪の表明や集団殺害・テロリズムを呼びかける憎悪表明が反差別の理念に抵触する点についても、国際的にも幅広い合意形成と勧告等の実践が蓄積されてきた。これらの事実は確固として存在する¹⁸。

この一方で、具体的になされた個別の発話について、「何が規制に服すべきヘイトスピーチに該当し、何がそうでないか」を識別・判断する基準については合意形成がなされておらず、またその規律も様々となっている¹⁹。

最も著名な具体例としては、ホロコーストの歴史を持つドイツ刑法における民衆扇動罪（第 130 条 1 項）があげられるだろうが、オーストラリアの 1975 年連邦人種差別法 18C 条の上に導入されたオーストリア刑法（1995 年）における集団に対する暴力の勧奨（第 80.2A 条）、その成員に対する暴力の勧奨（第 80.2B 条）、ジェノサイドの唱道（第 80.2D 条）等への刑罰付与も特徴的である。具体的には、後者では「人種、宗教、国籍、国民的若しくは民族的起源又は政治的意見」によって識別される集団に対して、故意に威力・暴力の勧奨を行うことが違法とされるが、この規律は専らオーストラリアの社会状況と白豪主義という歴史的背景に基づくものとされ、原則として発話が構成・発生させる具体的な害の発

¹⁸ いわゆる国際人権規約 B 規約（1966 年採択）及び人種差別撤廃条約（1965 年採択）には「ヘイトスピーチ」という語こそ表れないものの、差別的取扱いの禁止及び同取扱いを助長する表現についての禁止を、その制定時より盛り込んできた。例えば、国際人権規約 B 規約第 2 条第 1 項は「この規約の各締約国は、その領域内にあり、かつ、その管轄の下にあるすべての個人に対し、人種、皮膚の色、性、言語、宗教、政治的意見その他の意見、国民的若しくは社会的出身、財産、出生又は他の地位等によるいかなる差別もなしにこの規約において認められる権利を尊重し及び確保することを約束する」とし、表現の自由の一連の規定中における第 20 条第 2 項において「差別、敵意又は暴力の扇動となる国民的、人種的又は宗教的憎悪の唱道は、法律で禁止する」と定められる。後者の人種差別撤廃条約においては、その第 4 条において「締約国は、一の人種の優越性若しくは一の皮膚の色若しくは種族的出身の人の集団の優越性の思想若しくは理論に基づくあらゆる宣伝及び団体又は人種の憎悪及び人種差別（形態のいかんを問わない。）を正当化し若しくは助長することを企てるあらゆる宣伝及び団体を非難し、また、このような差別のあらゆる扇動又は行為を根絶することを目的とする迅速かつ積極的な措置をとることを約束する」との規定がなされている。その上で、国連人種差別撤廃委員会が、ヘイトスピーチに関する一般的勧告 No.35（2013）において、人種差別撤廃条約第 4 条が民族的・宗教的属性に基づくヘイトスピーチの禁止を含む趣旨であることについて明示的に言及している。文言を含む歴史的な経緯については、木下昌彦（2020）「Hate Speech in Japan を読む 国際法・憲法・比較憲法の観点から」に詳しい。

¹⁹ 国連人種差別撤廃委員会によるヘイトスピーチに関する一般的勧告 No.35 でも、犯罪として処罰することについては人種主義などの発言の最も深刻なものに留め、それに至らない発言などについては他の手段をとるように勧告されている点には留意せよ。

生で限定されるわけではない。さらにこの規律は、勸奨行為の時・場所・方法を問わない点から、社会的状況やオンライン空間の技術的環境の変化を通じ、広範な行為に及ぶこととなる²⁰。

これらと比較した際に、規制の緩やかな国として挙げられるのは米国や日本である。例えば本邦におけるヘイトスピーチ解消法（本邦外出身者に対する不当な差別的言動の解消に向けた取組の推進に関する法律：2016年）を見てみよう。そこでは「本邦外出身者に対する不当な差別的言動」を次の三要素、すなわち、（ア）専ら本邦の域外にある国若しくは地域の出身である者又はその子孫であって適法に居住するもの（以下この条において「本邦外出身者」という。）に対する差別的意識を助長し又は誘発する目的で、（イ）公然とその生命、身体、自由、名誉若しくは財産に危害を加える旨を告知し又は本邦外出身者を著しく侮蔑するなど、本邦の域外にある国又は地域の出身であることを理由として、（ウ）本邦外出身者を地域社会から排除することを煽動する不当な差別的言動をなす言動である、との要素から定義する。（ア）において多様なエスニックマイノリティーを対象として含まない過小包摂や、（イ）（ウ）における該当行為の文言上の曖昧さもさることながら、最も特徴的な点として、同法は制裁を含む執行発動根拠を持たないいわゆる理念法である点が挙げられる。すなわち、直接に集団侮蔑を法的な規律対象とするものではない。この点において、上記ドイツ刑法を筆頭とする諸法とは大きな距離を持つと指摘されることも多い²¹。

ただし、こうした不一致は単なるグローバルな規律形成の不徹底や怠慢ではない。（たとえ逡巡の末にせよ）そうならざるをえなかった、下記の理論的・実践的な難点があるためである。

（2）難点1 特定の属性と「それを持つとされる現実の社会集団」との間の距離

一方には、そもそものヘイトスピーチの定義をめぐる理論的難点がある。直観的に合意できる代表的な例をリスト的に並べたとしても必要十分な定義にはならない上、歴史的・社会的状況に鋭敏に反応するヘイトスピーチ現象について何らかの還元的な定義を求めることは（逆に定義から逸脱する例を示す反論を招くため）、規制可否という議論の焦点を逸することにつながるだろう。

第一に、**特定の属性と「それを持つとされる現実の社会集団」との間の距離**の問題がある。

²⁰ Sarah Sorial & Sam Shpall(2022)“”第6節では同法の運用実態に照らせば、規律が苛烈なものとなる場面は決して多いものではなく、主として法の表出効果に期待しているものであるだろうことを指摘している。しかし、こうした社会的状況やオンライン空間の技術的環境の変化を反映させない規律を下に、個々の表現の当否の判断を専らその時々の人権委員会の判断という手続過程や、オンライン空間であれば私的主体でありかつ媒介者であるところのプラットフォームを通じたコンテンツモデレーション、すなわち事実上の権力へと委ねるおそれを認識する必要があるだろう。この点につき、永石尚也（2019）「法の執行と正統性」『法政策の試み第19集』（信山社）の第2節を参照せよ。

²¹ Ayako Hatano(2021)“Hate Speech and International Law: The Internalisation of International Human Rights in Japan,” in (ed.) Shinji Higaki, Yuji Nasu, Hate speech in Japan : the possibility of a non-regulatory approach では、本文に述べたように本法のヘイトスピーチ解消法が文言上の刑罰の抑制性や非包摂性において人種差別撤廃条約の国内法化としては不十分である一方で、立法過程での広範なステークホルダーの関与や社会における議論喚起を通じて、立法後のヘイトデモ件数の急激な減少との相関を持つ点を指摘している。

「人種」や「皮膚の色」を筆頭として、過去に社会的に抑圧されてきた特定の集団的属性に関連づけずにヘイトスピーチがもたらす害を把握するのは困難である一方で、当該属性の多くは生物学的・解剖学的その他の明瞭な性質に規定された概念というよりは政治的論争の最中にある社会的構築物であり²²、マクロ的に変化に富む歴史的・社会的状況に相関的である。この結果、ある同じ名称で呼ばれる属性を（自己の一部として）持つとしても、それへの帰属やコミットメントの幅は激しく、同地域の同時代の内、さらには同世代の内であっても、制度ごと・社会ごとにその可視化や取扱い上の差異も濃淡を持たざるをえないだろう。

もちろん、過去時点の抑圧という一点において特定の属性を括ることはできるわけだが、その場合には当該カテゴリーは事実として属性を共有する具体的な集団を欠くことになる上に、また規範的にも（たとえ日常的には通用しつつも）誤ったカテゴライズを使用する負の効果として、既存の規範の強化及びそれによる周縁化を招く危険が付きまとう²³。逆にというべきか、それゆえにというべきか、例えば法においては（当該属性に基づいて別異の取り扱いをする傾向性を緩和・除去するために、当該属性をブラインドし、当該属性を理由とする取り扱い上の不平等を是正することを目指すというように）、当該誤ったカテゴリーについては専ら消極的に、いわば定義不能な属性として遇してきたとも言えよう。

ここで、その属性は専ら認識的なものであるとして、記述的にはある時点の・ある社会における（誤認を含む）認識により外延を記述する作業による解明で十分とされるかもしれない²⁴。ただし、その作業は概念の内包的な明晰化ではないことに留意する必要がある。仮に認識的に当該属性を理解する場合、誤ったカテゴライズを認識上は真なるカテゴライズと取り違える危険を冒すのみならず、カテゴライズ間の比較及び価値的な順位づけが困難

²² 前掲注)8 『バッド・ランゲージ』第5章[概念工学]を参照せよ。この上で、意味変化のコントロール可能性が問題となろうが、本稿では意味変化要因とコントロールとがマクロ的な歴史的・社会的状況に相関的であると考えており、この点につき意味変化要因の把握不能性とコントロール可能性の欠如を訴えるカペレンの「儉約フレームワーク」の説明（同書5-3-3-2）と本稿の立場とは軌を一にする。

²³ Judith Butlerに代表されるこうした本文記載の見解に対し、サリー・ハスランガーは規範的含意を受け入れた上で「抑圧の構造を強化するのではなく、むしろ切り崩す助けになることが見込まれている」ものとして、誤ったカテゴリーに対しても改良的なアプローチをとる優位性を主張していた。しかし、この「切り崩し」の有効性については、前掲注)に記載した「儉約フレームワーク」との関連で筆者としてはやや懐疑的である。サリー・ハスランガー（2022）「ジェンダーと人種——ジェンダーと人種とは何か？ 私たちはそれらが何であってほしいのか？」（慶應義塾大学出版会、木下頌子訳）を参照せよ。

²⁴ 荒井ひろみ、和泉悠、朱喜哲、仲宗根勝仁、谷中瞳（2021）「ソーシャルメディアにおけるヘイトスピーチ検出に向けた日本語データセット構築の試案」言語処理学会 第27回年次大会 発表論文集などを参照せよ。このように実際、コンテンツモデレーションの実施や判定基準の正当性確保に向けた調査として、多数の発話タイプを用意した上で、現にその言葉を見たときに直感的に「ヘイトスピーチ」に当たると考えるか否かを多数の被験者に判定させる調査がある。より具体的には、地域「独自の歴史・地域性を踏まえたヘイトスピーチ理解と、人文社会科学におけるヘイトスピーチに関する理論的知見に依拠したアノテーションガイドラインを設計」した上で、結果の統計的処理によって、当該地域、当該文化、当該時点、当該発信/受信メディアにおいて広く認識されるだろうヘイトスピーチ該当性とその濃淡を付与する、というものである。当該論文でも、語彙そのものが特定の文化カテゴリーと結びついている場合の偏見の混入などの限界が指摘されるどころであるが、より深刻なところとしてはアノテーターの属性を問わずにこの調査を行う点ではなかろうか。この場合には、特定マイノリティへの感受性や共感性の有無によるバイアスが混入するため、ロナルド・ドゥオーキンが『権利論』で提示した「外的選好」と同様の問題が生じることが懸念される。

に至るようになるために、新たなカテゴリーが提出されるごとにその「是正」を繰り返すこととなろう²⁵。また、そのカテゴリーの拡大傾向ゆえに²⁶、翻って規制態様は制限的なものに留めざるを得ないことにも帰結する。

さらに、当該属性と現実集団との間にある歴史的過程にも着目できる。過去から継承された様々な属性に付随した抑圧があることは事実としてある。他方、それら抑圧は各々が質的に異なるとともに、それらを支えた当時の経緯やイデオロギーもまた統一性を欠き、さらには過去のある時点から現在に至る回復や和解、その後の「赦し」の歴史的過程及びその実践も様々であることに留意する必要がある²⁷。こうした（ときに原理的な証言者の不在性²⁸と

²⁵ 安藤馨・大屋雄裕（2017）『法哲学と法哲学の対話』（有斐閣）第5章における「応答」末尾の問題提起を参照せよ。なお、「展望論考（近刊）」第4節で後述するように、ジェレミー・ウォルドロンがヘイトスピーチ規制における属性の無制限の拡張を問題視し、「アイデンティティ政治の無責任さ」と集会的決定の挫折（『ヘイトスピーチという危害』159-162頁）と評していたことが想起されるだろう。

²⁶ 例えば沖縄県における「沖縄県差別のない人権尊重社会づくり条例」（2023）においては「県民」という属性に着目して、その地域社会からの排除煽動などが生じた場合における言動解消措置を講じることを定めている。しかし、出でさえない「県民」性の範囲が問題になるとともに、（もちろんその歴史性に鑑み、他の県と異なるカテゴリーの特殊性は認めつつも）「県民」と制度的な地方自治体という公的・政治的アクター性との密接不可分の結びつきによって、有用な表現・議論をも縮小させかねない懸念についていかなる対処がなされているのかは、以後3年ごとの見直し運用において必然的に争点となるだろう。

²⁷ 林志弦（2022）『犠牲者意識ナショナリズム』（東洋経済新報社・澤田克己 訳）の言葉を借りれば、グローバルな「記憶の戦争」の問題が表面化しているとも言えるだろう。ホロコーストが人類史的な災禍として認識されている事実には疑いはないが、一方では「ホロコーストのコスモポリタン化」現象と呼ばれる脱地理化のねじれがあり、他方では他の民族的・宗教的虐殺がそうみなされているわけではない（同時代の「西側」を含む各国において、身体障害・精神障害を持つ者に対する虐殺や隔離がこれと並行していたことを思い起こそう）緊張関係を抱えている。このように何が人類史的な災禍と見做されるかをめぐる林日くの「記憶のポリティクス」がここにあらわれる。

また国際人権規約B規約や差別撤廃条約の文言修正の経緯において、ヘイトスピーチに対する刑罰化を強く主張したのが専ら「東側」諸国であり、「西側」諸国が（憎悪言動の唱導の排除の重要性に鑑みても）制裁を背景とする禁止がもたらす負の効果の観点から例外を設けさせた経緯を、ここで思い起こすことができる。この経緯は、各人権類型の間の緊張関係を内包する統合的な人権保障の文脈においても、「記憶のポリティクス」の動きとそれが規定になった後に普遍的なものとして通用してしまう事実を取り込む必要性が明らかであることを示す範例となる。この経緯につき、前掲注）木下昌彦（2020）91頁以降を参照せよ。

²⁸ ジャック・デリダ曰く「生き延び」と「相続」の問題がここにあらわれる。本稿の主題の中心からは外れるため以下簡潔に要点のみ記す。「記憶の政治」におけるホロコーストの継承者は（第二次世界大戦を横断する反共のうねりの中、複雑な経路を経て）終局的にはイスラエルと目されるに至った。しかし、第一に事実として生きている者はホロコーストの害を直接に受けたものではなく、また生き延びた目撃証言者等にとってもトラウマティックな苛烈性のため、証言・叙述の真実性と歴史的な真実との関係をどのように捉えるのかという課題、すなわちホロコーストの「表象不可能性」をめぐる課題が現れた（1990年の「〈最終解決〉と表象の限界」会議の論争当事者の一人、ヘイドン・ホワイトによる『実用的な過去』（2017、岩波書店、上村忠雄 訳）、特にそこに取められている「真実と環境」（2012）及び「歴史的な真実・違和、不信」（2016）に詳しい）。第二に、より政治的な課題も含まれる。つまり、上記の表象不可能性にもかかわらず「ホロコーストの子供たち」の地位がイスラエルという国において偶然的に分有され、系譜＝家族史とは関係なく、社会的記憶の領域で特権化されるにいたった歴史的偶然性が厳然としてある事実である（この点は、前掲注）27 林志弦『犠牲者意識ナショナリズム』（2022年）の第6章の[世襲的犠牲者意識とイスラエル]の項（225頁以下）にも詳しい。）。この振れをまずは見てとることは、ホロコーストの害を矮小化することなくとも指摘できよう。

より実践的な観点から見ても、個別の被害者の（例えば証言内容を公にすると侮蔑的な視線に晒されるなどの故の）語りの脆弱性あるいは「証言の飲み込み testimony smothering」と、その（公のアーカイブにする営為から離れれば忘却されるが故の）語りの不在・消滅に、「記憶」をめぐるポリティクスは挟撃されているのである。「証言の飲み込み」については、クリスティ・ドットソン（2022）「認識的暴力を突き止め、声を封殺する実践を突き止める」『分析フェミニズム基本論文集』（慶應義塾大学出版会、木下・渡辺・飯塚・小草 訳）の特に217頁以下を参照せよ。

も呼ばれる) 事情のため、直観的に把握しやすい既存の被差別的属性の選言的列挙は、現在時点において保護されるニーズを示唆する政治的無力性の有無・程度と当該語彙で把捉される現実の社会集団との関連性を欠く問題を、往々にして生じさせることとなる。

結果として、その時々において政治的に無力化された社会集団を過不足なく取り出し、そのニーズを満たすために「ヘイトスピーチ」という語の分析から発して有効に対象を区切ることは困難に至るだろう。

(3) 難点2 特定の属性と「それを介した人格形成及び人格毀損という出来事」との間の距離

第二に、**特定の属性と「それを介した人格形成及び人格毀損という出来事」との間の距離**の問題がある。過去に社会的に抑圧されてきた特定の集団的属性は、ミクロ的に個人の内部において人格の一部に組み込まれ不可分となることもあるが、他方で当人の全人格は当該属性に尽きるものではなく、複数の属性の複雑な束ね上げによって構成される。すなわち、抑圧の源泉はある一つの属性に限定されないとともに、逆にある一つの属性を持っていることが抑圧されていることを意味するものではないのである²⁹。集団と個人とは、ある時点におけるある属性の共有や当該属性への社会的視線によるカテゴライズにおいて包摂関係にあるとともに、その包摂関係もまた時間の中で揺らぎ、離脱や新規参入、あるいは社会的な呈示の過程で変容する以上、(法的紛争事例としての団体の内部関係を取り上げるまでもなく) 一定の緊張関係に立たざるをえないだろう³⁰。

もちろん、当該属性が当人の良心や信条といった深く人格を規定する領域に根ざしていたり、人生選択に影響を与える変更困難な特徴(例えば心的外傷を含む不可逆かつ深刻な被害などの出来事そのもの)を証跡する性質³¹等であれば、事実的にも規範的にも人格と当該属性との結びつきは異論なく認められよう。しかし、そこにおいても何が人格に深く根ざしているのか、あるいは不可逆に当たるか、さらに誰がその被害を共有する者たちの範囲に属するとされるのかは、多くの場合論争的であらざるをえないことも事実である³²。まして

²⁹ アイリス・マリオン・ヤング(2020)『正義と差異の政治』(法政大学出版局) 58頁

³⁰ 松宮孝明(2018)『刑法各論講義』など、法人への侮蔑罪について、法人に属する構成員への侮蔑となる場合と、単に法人のみへの侮蔑にあたるものとで成立可否を分ける見解が、ここでは想起されよう。

³¹ 前掲注)8『バッド・ランゲージ』ではまさに「レイプ」概念の概念工学性が論じられている。関連はするもののより複雑な社会構造を背景とする課題として、朴裕河『歴史と向き合う』(2022年)の第3章で取り上げられている「慰安婦」の国境による分断事例が挙げられる。本稿と関連する限りで要約すれば、同書ではまさに具体的被害と直結した「慰安婦」の問題が、(歴史的に偶然的なものでありながらも2000年の女性国際戦犯法廷に至る過程で)「戦争犯罪」の枠組みで理解されていく中で、被害者の属性もまた「交戦国」か否かという図式の下に包摂されていくこととなった問題が指摘される。すなわち、「戦争犯罪」という語りは、同時にこの問題を帝国/植民地問題として語る語り口を阻むこととなったことがある。この結果として帝国/植民地に属していた個々の被害者同士(朝鮮人・台湾人・日本人を含む)は分断され、「語り」の紐帯と連帯とを失わせることとなる、というのである。無論、この問題は「慰安婦」問題に限定されることはない。責任帰属をめぐる議論は一定の落ち着きを見せつつあるがいわゆる「徴用工」問題でも、この構図が、臣民動員の旗印の下、実態としては社会における低階層や家庭内でも次男以下など「弱者」徴用=階級動員として反復されることとなったものであろう。

³² 前掲注)27 林志弦『犠牲者意識ナショナリズム』(2022年)第5章の[追放・抑留と加害の忘却]の項(173頁以下)

そうではない属性であれば、たとえ引き剥がしがたい性質を持つとしても、少なくとも複数の集団的属性が重疊的に作用するのが価値の多元性を標榜する現代社会においては通常であるものとして取り扱われるかもしれない。当該属性を取り巻く社会的階層と行動選択の可能性や趣味選択（その表明）の可能性などが強固に結びつくことで実現困難となるなど、極端な抑圧的な社会構造を持つ状況においては別であろうが、それは属性そのものの問題ではなく、それとの結びつきに由来する問題である。

翻ってオンライン空間を介した「人格」は、後述するようにその形成段階及び（選択的）呈示段階において、リアル空間よりもその場の規範にフィットする「顔（=面目）face」として操作可能（fabricable）性を介しているとの側面を持ち、またその側面の利益を有していることもまた否定し難い³³。これらの結果、ときに混同されがちだが、当該属性を介したヘイトクライムや差別的取扱いが当人に直接に届くのととは質的に異なり、ある属性を介したステレオタイプの表現それ自体による当人の人格性の毀損という害は間接的にのみ当人に届くことを前提に、規制を組み立てる必要が生じるのである。

（４）発見的用法としての「ヘイトスピーチ」？

で詳述されている。それによれば、例えばドイツでは、ホロコーストの加害者としての自己イメージが広く共有される一方で、多くのドイツ人こそナチの第一の犠牲者であるのだとする捉え方も根強く、とりわけ東欧からの「被追放者」としての集団（1200万人が追放され、うち50万人から200万人が引揚げ時に死亡したとされる。）にはその苦難の歴史をホロコーストに連なるものとして持つ者たちも少なくないとされる。事実、追放を受けたドイツ人に対する（ポーランドおよびチェコにおける）抑留収容・虐殺の悲劇は、加害者集団に属するとされた無名の者たちに対する暴力の無際限さを可視化するのであり、上記見解も故なきものではないかもしれない。日本における対応物としては、自分たちを軍部に騙された犠牲者だとする認識と、GHQのオリエンタリズム（＝「日本国民は権威に服従する封建的な慣習の奴隷」であるとする考え）との共犯関係を筆頭に、関東大空襲をはじめとする民間人への空爆や二つの原爆投下による犠牲（そしてシベリア抑留）を頂点としてホロコーストの悲劇と重ねられて理解されてきたことが挙げられようか。これについては、ドイツでは早々に絶版となっていたヴィクトール・フランクルの『夜と霧』が日本語翻訳でのヒットを経て、世界的なベストセラー化を果たしたという過程も目に付くところである。現在、「ホロコースト」が虐殺一般を指すようになりつつある「グローバルな記憶」の濫用を避け、かの一回的な出来事としてのユダヤ人虐殺については「ショア」という別名を用いる傾向が強まっているのも無関係ではなかろう。以上が林の見立ての概要である。

なお、この点に関し、福間良明「断絶」の風化と脱歴史化 メディア文化における「継承」の欲望（2016）が参考になるとおもに、同氏による「記憶の「封印」と「発明」--「継承」と「忘却」の歴史社会学」（2016）に挙げられている知覧特攻平和会館（1975年設立）とその後の「語り口」の変化にも、同様の構造が見てとれよう。同論考では特攻隊基地を有していた沖縄・知覧町における特攻隊員らについての「語り」の持つ歴史的な「歪み」が検証されている。そもそも特攻隊員らは全国から集められたのであり、知覧町の外部から来た者たち、すなわち知覧町民とは系譜的な関わりは全くない。実際、この外部者によって知覧町の主要産業である茶畑は接収され、経済的にも打撃を受けたとされる。加えてその後、終戦直前における激しい沖縄戦（及び他ならぬ日本軍の手による惨殺）の記憶からすれば、単なる外部者というよりも更に敵対的な位置を、知覧町民の記憶に刻んでもおかしくなかった。しかし終戦後、占領下における本土復帰要求の高まりにあっては、特攻隊員たちを含め、自分たちの内部に属するものとして遇される。その後、特攻隊基地跡地には知覧特攻平和会館（1975年）が作られることになるが、それは公的には、特攻隊の記憶が知覧町住民固有の「語り」に取って代わった瞬間でもある。嘔み潰された、屈託ある、有象無象の情念が、そこでは覆い隠される。公的に他者の記憶を継承することで、いわば自らの記憶を積極的に忘却へと追いやってしまうのである。知覧の記念館は、その歴史と地理の地場から畏怖（イフ）を消し去り、「封印」とともに、公的記憶を発明しおおせてしまった。以上が福間の見立てであり、本稿との関連を有する。

³³ Th. Nagel(1998), "Concealment and Exposure", *Philosophy & Public Affairs*, 27 (1) 及び同論考への解説を含む永石尚也 (2020)「プライバシー・監視・アーキテクチャ：「AI」と法」の余白『法政策学の試み 第20集』（信山社）第2節（特に「顔」についての言及は129頁）を参照せよ。

これら2つの難点から、特定の属性に照らしたヘイトスピーチの定義そのものを放棄する誘惑に駆られるかもしれない。たとえば、属性に基づく差別的雰囲気やの瀾漫という事態をもたらす発話一般を「ヘイトスピーチ」とする、あるいは社会に内在する差別的構造を発見するための議論喚起力のある言葉遣いとして用いることを許容する、といった「解決」策である。明確な差別的社会的構造の存在があるならば、表現はその発現そのものであり、被差別集団への危害の告知に類した作用を果たすことになるはずだ、とそこでは主張されることになる。

しかし、この捉え方は両義的である。前者について言えば、そもそもどの属性が差別と関連するか、そして発話が差別的雰囲気を瀾漫させていることをどのように認定するかを恣意化させることは、逆にその後の発話一般への介入を放縦化させる事態を招くこととなる。さらに、当該発話と名誉を毀損する事態を結びつけることは、むしろ発話の社会的意味を、固定・強化してしまうことになる。例えば、ヘイトスピーチに属する発話がなされた際、専ら発話者の社会的評価が毀損される社会となるか、それとも発話を受けた者の社会的評価（あるいは名誉感情）が毀損される社会となるかは、現実の社会の状況に照らせば、少なくともその時点では多くの社会が不確定であらざるをえない³⁴。それにもかかわらず、前者を望ましいものとして確定的に選択することは、当該発話そのものをスティグマ付与的なものに固定・強化する点において、被害者救済にむしろ反しかねない。この意味において問題である。そもそも、事実が不確定な状況において、実証的知見の蓄積と適用に際しての修正過程を経ることなく害の確定性を選択できるとする考えは、極めて強い特定の立場にコミットメントすることを前提とすることに自覚的である必要はあろう。

後者についても同様である。差別的構造を発見し、告発するために「ヘイトスピーチ」という語を用いるとすれば、翻ってその語自体が自分に対しても望ましくない発話や嫌悪をもよおす発話をラベルづけするための語として誤導的に使用されうることをも許容することを意味する。さらに、そうした慣行が続けば、適切に発見・告発した状況においても、誤導的に使用しているのだらうと軽んじられる事態を自招する結果にもなるだろう。例えばこのような誤用の例としては、2022年に始まるウクライナ・ロシア戦争初期において、「ウクライナ政府によって虐待されてきた人々を保護する」³⁵ため、すなわちウクライナによる国家的ヘイトクライムを告発することで成り立つ名目が、ロシアによる侵攻正当化の理由として掲げられていたことが記憶に新しい。こうした名目的使用を許す場合、「虐待」という語彙を用いて差別的構造を発見し、構造を是正する機能は弱化させられることとなり、特に国家作用としての規制実践へと結びつける根拠としての適格を欠くに至るだろう。「ヘイ

³⁴ プライバシー侵害をなす発話などのように、構造的な不可逆性が顕著であるものは別に論じうる。しかし、プライバシーについてできても、何が知られることを望まないプライバシーに属する事項であるかは当該社会の状況に即して流動的であり、また社会において列遇されている者の場合、別の経済的・社会的利益との関連でプライバシーを自ら放棄させる状況をも存在するために、その確保が一枚岩ではないことは言うまでもない。前掲注)33 永石「プライバシー・監視・アーキテクチャ」(2020) 第2節・第3節を参照せよ。

³⁵ 各種報道に表れている2022年2月24日のプーチンによる軍事作戦開始演説より引用した。

トスピーチ」についても類似の議論が当てはまる。ホロコーストを典型とする地球規模の記憶文化は、「遠く離れた他者の痛みに共感する道徳的感受性を強める契機」となりつつも、「集合的記憶の民族主義的な競争を地球規模に拡大し、深刻化させる」³⁶と指摘されることもあるが、まさにホロコーストの裏面であるヘイトスピーチもまた上記と相同的な構造を持つ。属性に基づく差別的雰囲気やノイズをヘイトスピーチとみなす「解決」とは、常にこの不都合なトレードオフに陥る危険と隣り合わせなのである。

以上の検討から明らかになるように、規制文脈で求められているのはヘイトスピーチの**還元的な定義**をめぐる探求（やその放棄）というよりも、規制に服すべきヘイトスピーチとそうでないヘイトスピーチの**概念的な区別の改訂実践**である。この改訂実践のためには、上記のように特定の属性とそれへの差別的な社会構造から発してヘイトスピーチの対象を特定する困難な経験的検証を続けるとともに、ヘイトスピーチと目される発話の言語としての性質を概念的に明らかにする必要があるだろう。具体的には、ヘイトスピーチと目される発話の（専ら発話者側の）性質のみならず、ヘイトスピーチの害の（発話者、聴衆、発話のターゲット集団、社会構成員等を横断して生じさせる）構成／発生機序を、他ならぬオンラインヘイトスピーチに即して正確に把握する必要がある³⁷。

1-3. 規制をめぐる困難② 発話上の特性

（1）難点3 規制対象を画する実践的難点

さて他方で、（属性ではなく）発話の言語的特性に着目して規制に服すべきヘイトスピーチの限界を画することも実践的な難点を抱えている。より詳細には論考[展開]にて詳述しているため、ここでは発話内・発話媒介的アプローチ双方を横断して生じる課題を、簡潔に指摘するに留める。

まず、ヘイトスピーチの外延をその発話内容のタイプから特定することには困難を伴うことが知られている。「憎悪 hate」という発信側の事情に主として着目するかの文字列上の

³⁶ 前掲注) 27 林の補論部分 (362 頁以下) を参照せよ。

³⁷ Erving Goffman(1981), *Forms of Talk* では、対面的な相互行為の分析において、発話と受話（参与の身分 participation status）との関連が整理されている。発話産出段階においては、発話主であるところの animator、発話内容の意味・表現を選択した author、発話帰属点としての principle、言及において登場する figure というフォーマットが前置される。受話（参与の身分 participation status）段階においては、そのバリエーションとして、直接の聞き手である recipient、発話が向けられた対象である addressee、集まりへの参加が公式に認められた ratified participant と認められていない漏れ聞きを意図して行う eavesdropper、たまたま聞く機会を得てしまった overhearer とそのうち発話者側から把握可能な bystander といった身分が列挙されている。これらが会話の中で前提とされるとともに、続く非難や弁明の中でそれを改訂していくことで会話において適切な発話の条件を更新していくという相互行為の描像は、David Lewis が 1979 年に "Scorekeeping in Language Game" において提示した会話の Scorekeeping のアイデアにも適合するものであるものといえよう。なお、Goffman はこの過程が単に会話の適切性を超えて、参与者たちのアイデンティティや続く行為、文化へと接続されることを想定して分析を行っている点において、本文に述べた本稿の関心に沿う。詳細は論考[展開]に譲る。

見かけとは異なり、ヘイトスピーチは、表現内容を分析することで辞書的に（選言的に拡張したものであれ）定義される類の、特定の発話カテゴリーに属する発話集合としては把握されないのである。

例えば、これを蔑称や隠語と比較してみよう。蔑称そのものは特定の記述の対象を指示するものではなく、さらには侮蔑的な発信者の意図を前提とするものでさえもないことから、蔑称の発話による害は、蔑称の持つ意味内容を分析することから導き出されるのではなく、その受信者に対する社会的な効果（典型的には害を根拠とする禁止と禁止違反に対する諸制裁）から遡って把握される他はない³⁸。語彙そのものの効果は、社会的な効果を支える文脈依存的である。ヘイトスピーチもこの点においては同様である。憎悪や害意なしにもヘイトスピーチは発話しうるわけだが、それはヘイトスピーチの持つ意味内容からではなく、その受信者に対する社会的な効果から遡って把握されることを示唆する。すなわち、何が社会的に害を構成し、あるいは発生させるかは、当該発話が置かれた歴史的背景・社会的状況を踏まえて把握されるだろう、専ら現在の差別的慣行との関連性を有する広範な文脈に依存する。

この広範な文脈依存性から見えてくるのは、特定の発話内容を典型的に把握し、地理的・文化的に多様である現在の差別的慣行と個々に結びつける作業の経験的な困難性である。例えば、ある属性（例えば性別）に着目した蔑視的な表現が、当該集団に属する者一般を「格下げ subordinate」し、あるいは「構成する construct」する効力を持つと主張されることがある。しかし、集団と個人との関係には、部分的なものから全面的なもの、着脱可能なものから不能なもの、可視的なものから不可視的なものまで、さまざまな場合が想定される。この多様性を無視して一律に効力を想定することはできない以上は、この効力を発話媒介行為として見た場合には、社会状況と個別的場面に依存したその実現過程の実証的裏付けに依存することとなる。そのため、その妥当性の検証なしには効力の有無・程度は判定不能となろう。あるいはこれを発話内行為として見た場合には、果たして発話者の権威や状況の支えによることなく効力を生じるのか、さらには一様にその効果を生み出すのかとの疑いが出されることとなる³⁹。

次に、発話のフレーム依存性が挙げられる。（受信者の訴えが認識上の端緒となりうることは当然としても）受信者に対する社会的な効果と言っても、個々の受信者にとって「そう

³⁸ 前掲注)8 『バッドランゲージ』第6章[6-5節] 159頁における蔑称の「禁止説」を参照せよ。なお禁止説の難点として、害に対する禁止の理論的先行という理論上の難点が挙げられる。しかし、この指摘に対しては同説を一般化し、ネガティブな効果を生み出す歴史を持つ動的性質を持つものとして捉えることで解消される。さらに、後述するように法の作用は、この動的性質の上で禁止の限定（単なる悪い表現と禁止されるべき蔑称との区別）と制裁の限定（蔑称の中で、制裁が課されるべきものとそうでないものとの区別）を理論化する営みである以上、本稿の以後の展開上は問題とならない。

³⁹ 江口聡（2022）「分析フェミニスト哲学者によるポルノグラフィ批評（悪いポルノ、悪い哲学）」応用哲学会第14回年次研究大会報告資料を参照せよ。本文で述べた蔑称使用の受け取られ方に類似した議論として、性的に露骨な表現（ポルノグラフィ）を持つ作品鑑賞の経験が挙げられる。同ジャンルに属する作品は、特に性差別的・暴力的なそれであってさえも（性差を問わず）一方的な被抑圧的立場を投影することなく鑑賞する態度を、比較として検討の俎上に挙げられよう。

聞こえた」ことや「そう読めた」ことからこの効果を直接に把握することはできない点が挙げられる。ましてや発話が特定の「メディア」を介してなされる場合には、上記文脈を形成するそのメディア特性が持つ効果も無視することはできない。言語や行動を成り立たせ、解釈させるフレームが重要となるのはここにおいてである。

例えば、ラジオでのフィクション作品の朗読劇は、たとえ一連の発話が報道調でなされ、たとえそのように広く誤解されようとも報道（誤報）とはならないし、あるいはフィクションならざるテレビのコメディアンが発話がいかにこちら側に語りかけているように思えてもそれはフレームが共有された会話ではない⁴⁰。テレビというメディアを介したそれらの発話は、画面内出来事のフレームの下にあり、画面を跨いで存在するのではない。

もちろん、こうした先行フレームを誤解する、あるいはその複層性を特定フレームに引き寄せて誤用する事態は、物理メディア（銅像等の象徴的建造物等）⁴¹、紙メディアにおいても起こってきたかもしれない⁴²。とはいえ、そうだとすると、漫画作品の特定のコマ中に現

⁴⁰ Goffman(1981), *Forms of Talk* で提示された例であり、実際にラジオを通じて集団パニック現象を引き起こした「宇宙戦争」事件などの事例を基礎としている。劇中に登場した蔑称は、フィクション内において蔑称として機能するし、フィクションを理解する上でそう機能しなくてはならないが、現実のあなたに対する蔑称使用ではない。テレビのコメディアンが用いた蔑称は有限の電波を用いて流すべきではない蔑称使用とはなりえても、あなたへの蔑称使用ではない。ましてや、画面のこちら側のお茶の間からテレビ画面の中のコメンテーターにツッコミをいれたところで、せいぜいお茶の間にいる別の人のフレームにおいてのみ効力を発する（もしも部屋に一人ならば孤独な擬似的会話に終わる）にすぎない。

⁴¹ 大屋雄裕 (2020)「赦しと忘却」アステイオン 93号 110-125 頁を参照せよ。そこで大屋は、キャンセルカルチャーの一つと目されるリンカーン像をめぐるある騒動に、物理的な抹消を介した記憶と赦しのダイナミズムを完全に失った「責任なき永遠の静寂」を見ている。「ワシントンDCにある解放記念像——白人大統領であるリンカーンの足下に黒人奴隷が跪いている彫像をめぐる騒動をとらえた動画も、議論のさなかで注目を集めていた (https://twitter.com/_sagnikbasu/status/1276682947895197697 以下のツイート)。「なぜそんな像を守るのか」と金切り声で叫ぶ女性のBLM運動家に対し、初老にさしかかった頃だろうか、向き合っている黒人男性が静かに問いかけを繰り返している——「誰が払ったのかね」「誰があの像のために資金を出したのだね」。よく知られているとおり、この像が描くのは南軍の首都リッチモンドの陥落後、同市を訪れたリンカーンを奴隷解放に感謝するために取り囲んだ黒人たちの一人が跪いたところ、「私ではなく神に対してのみ跪きなさい」と説いて立ち上がらせたエピソードである。この彫像もまた、解放された元奴隷たちによる献金で築かれたという過去を持つ。歴史的文脈のなかに置けばむしろ人種解放と平等の象徴であるものを、独立の存在として瞬間的に見ることにより、まったく逆の意味のもとで理解してしまったということになるのではないだろうか。このエピソードは、キャンセルカルチャーの基礎と帰結を象徴的に示しているだろう。すなわちその基礎は記憶の破壊と無知であり、帰結は無反省な自己中心主義である。たとえばあるシンボルが文脈抜きに見られた場合に一定の危険性を孕んでいるとしても（ジェファーソンやグラントはその例かもしれない）、その対策としてはBLMが目指すような情報を減らす解決ないし破壊ではなく、情報を増やす解決——そのシンボルが持つ意味を歴史的文脈のなかで示し我々の反省を促すような注記を添えることが考えられるはずだ。それはまた、情報が適切に与えられれば人々（の多く）はその適切な評価が行なえるだろうという人民への信頼に基づいている。それを拒絶するキャンセルカルチャーは自己の判断力を人民一般に対する高みに置いていることになるだろうがその足下を支えるものはきわめて怪しく、だからこそとすべきか、自己の信念に背く可能性のあるもの、反省を強いる可能性があるものを根こそぎに破壊することが解決策として模索されるのだ」。本稿では末尾においてこの論点をあらためて取り上げる。

⁴² 折しも2023年初頭、広島市教育委員会による教育教材から削除されたことで話題となった『はだしのゲン』は、（削除の可否自体が法（哲）学的にも興味深い話題を多数含むものの）まづもってそのメディア特性として商業漫画作品の形式をとる点に着目できる。教育目的の教材と商業フィクション漫画とは一般的に区別され、教育現場・図書館実務では後者を基本的には排除してきた中で、『はだしのゲン』がある意味で「特権的」に配架・使用が許容されてきた経緯は、近時の同書の教育過程からの排除の可否を議論する上でも前提とはなろう。その上で、同作の教育教材からの削除措置に対する肯定的・否定的な反応において（動機はともあれ）焦点化された対象が、方や反戦漫画としての位置づけから歴史の忘却につながるために削除はすべきでないとし、方や専ら戯画的・誇張的な表現方法や侮蔑的・差別的

れる登場人物による蔑称や差別的言動は、(漫画の読み方に通じていない、あるいは当該作品が置かれた歴史的な脈、当該ジャンルに通じていない読者にとっての語彙への直感的な反応を除いては) 物語全体における当該登場人物の性質を表示する(その必然性から要請される)ものであり、現実における蔑称としては一般的には機能しない⁴³。すなわち、物語内における蔑称・差別的言動としてまずは機能した上で、次いで物語全体の中でそうした発話への評価が解釈され、位置付けられることとなろう⁴⁴。登場人物が語りかけている対象は、文字通りには現実のあなたではなく、その発話の害は物語において閉じるフレームの力学とともにある。

同様に、こうした事態はオンライン空間においてさらに増幅する。孤独な検索行為に勤む際、原理的にはあなたの活動は追跡可能であるが、あなたの経験的には窃視的である錯覚を与えるかもしれない。SNSで「要領を得ない」意見を見かけても、それはあなた宛のものではない。仮に引用リツイートやリプライを受けた際にでさえ、そうでありうるのである。これらは、ブログやSNS等の発信・受信の形態と使用のされ方が構築してきた諸フレームの混線が起きていることを示唆するだろう。このことからすれば、オンライン空間で蔑称が使われる際にも、その語彙によって生じる効果はフレームの誤解によるものであることが大かかもしれない。そうだとすれば、蔑称や侮蔑的・差別的発話を専らその語彙(効果)にのみ着目し、媒介するメディア効果抜きに抽象化して論じることは、当該発話の機能を見誤ることにつながるだろう。とりわけオンライン上のヘイトスピーチについては、後述する拡散性や永続性という性質から、より広範な「害をなす」発話とともに、それへの自動検知やプラットフォームによるコンテンツモデレーション等、手法面・主体面での拡張もなされる。これは同時に従来の理解を主体面・作用面において拡張した準-規制による新たなリスクをも生じさせるだろう。複数の「メディア」を介し、法規制もあるいは社会的非難を含む事実上の規制作用もなされるのが通常であるオンライン空間においては、上記の文脈読解にあたり、当該メディアの技術的・社会的特性を踏まえた検討が要されるのである。

(3) 規制対象を画する実践的難点の上で

上述してきたように、オンライン上のヘイトスピーチに発語媒介的アプローチおよび発語内的アプローチを適用する段においてはずれが生じる。この理由は、原理的に言えば、オ

表現あるいは「ヒロシマ」差別への接続の懸念を問題視し削除もやむなしとするなど、論点及びその評価においてすれ違いを見せていた点は、こうした「誤読」の例として興味深いように思われる。

⁴³ 戸田山和久(2016)『恐怖の哲学 ホラーで人間を読む』NHK出版を参照せよ。例えば娯楽作品としてのホラー映画作品を鑑賞する際には、まさに恐怖感を求めて楽しむわけだが、ここでの「恐怖感」の意味が現実の「危機感」などとは質的に異なるのはなぜかと美学的・認知心理学的な議論から問うことができる。フィクションを介して「恐怖感」を避けようのない(ある意味では自分に責任のない事態)として安全に享受することは、ある感覚や行動に通じるシミュレーションでもある。その多層化したメカニズム自体が、ポルノグラフィを筆頭として議論される、いわゆる「低価値表現」一般との関係でも、単純化を許さない構造把握を要請するものと思われる。

⁴⁴ デイヴィッド・ルイス(1995)「フィクションにおける真理」(樋口えり子訳)における議論を参照せよ。なお、関連して野上志学『デイヴィッド・ルイスの哲学 ーなぜ世界は複数存在するのか』(青土社)が参考となる。

ースティンが想定していたような発話（の時点・文脈等）と発話内行為及び発語媒介行為との同時性が成り立たないことに由来するだろう。

これら難点を踏まえれば、ヘイトスピーチとは、当該発話がなされた歴史的背景・社会環境において、ターゲットとして名指された集団に属する個人に特殊な害をなすものとして、**機能的に把握された発話集合**として、害を発生させる個別場面ごとの分割と、その不断の文脈改訂に即応した改訂実践の中で措定される他はない。ここでの害の候補としては、身体・生命への危惧感や名誉感情・社会的評判の毀損、民主的熟議の基盤の破壊、あるいは多義的であれ「尊厳」の毀損などが挙げられる。ただし、上記のように機能的に把握される特性から、同じ発話内容であっても個別の社会環境によって発生する害の種別やその発生機序は異なり、また集団の種別により個人のアイデンティティに結びつく態様もさまざまに異なることは、規制の適用にあたっての前提をなす。さらに、規制自体の性質としても、時間を通じて（一定の解釈による幅を持ちつつも）維持される固定性を発揮するために、文脈上の変化に富んだ状況においては過小・過剰な規制作用を発揮する傾向性を持つ。こうした機能上のグラデーションを把握することなく、発話内容そのものの持つ害と発話が構成・発生させる害とを取り違えることで、容易に規制は文脈に適合しない誤作動へと通じてしまうだろう⁴⁵。

ここからの教訓として、ヘイトスピーチを広く法規制の観点から論じる際には、当該ヘイトスピーチが置かれた諸文脈を形成する要素を特定し、それら諸要素の結びつきから害の構成／発生機序を明らかにした上で、（効果の維持・消失過程をも含む）時間を通じたコントロール手法を提示する必要がある⁴⁶。オンラインヘイトスピーチ特有の害の機序に対応した、多元的な規制枠組み・非規制枠組みとしてこれらを提示する必要があるのである。

⁴⁵ ある規制が対象とした（例えば反イスラムの）ヘイトスピーチへの規制によって、別の保護されるべき集団の自由を逆に制約したり、まさに保護されるべき集団への別の害を生み出すことに留意が必要であることは、前掲注)19 で人種差別撤廃委員会も注記している通りである。

⁴⁶ 齊藤愛（2015）『異質性社会における「個人」の尊重』。いわゆる表現の自由の保護範囲論の枠内で集団侮蔑を内容とする表現の使用を論じる憲法学上の議論もまた、上記文脈を構成する一要素として包含される。別稿の論考[展開]第4節において詳述する。

2、議論の前提：法的介入と各種のリスクトレードオフ

---R. Simpson による J. Waldron 批判の検討を通じて

2-1. ヘイトスピーチの害の構成/発生機序と「安心」の毀損、そのコミュニケーション形式・空間特性への依存性

(1) コミュニケーション形式の特殊性

さて、SNS やオンライン・フォーラム等、オンライン上においても無論、前述の機能的な意味で把握されたヘイトスピーチに類した発話はなされうる。そのうちのいわゆる「特定」やつきまといなどリアル空間における行動や被害に繋がるいくつかのものは、従前から存在した（対面状況で畏怖を与え、静穏な生活を害する⁴⁷類の）リアル空間における示威的なヘイトスピーチに近い性質を持つし、また著名人による SNS での継続的発信等いくつかのものは、雑誌やテレビといったマスメディア上における従前から存在した差別助長的なヘイトスピーチに近い性質を持つだろう。

事実、（民事賠償はもとよりとして）明示的に特定個人に宛てられた前者に属するヘイトスピーチ群については、本邦でも脅迫罪・侮蔑罪等に当たる行為である以上、実際に刑事罰が科されてきている⁴⁸。こうした事例において、オンラインを通じた個人に対する差別的・侮蔑的発話が、少なくとも報道に現れる段階で特段の断りなく「ヘイトスピーチ」事例として認識されているのは、こうした事情によるものである。

しかし以下で詳述するように、後者に属するオンライン上におけるヘイトスピーチ群の多くは、リアル空間の対面状況においてこれまで蓄積されてきた事例群⁴⁹とは性質を異にする。それらは、技術的には双方向的でありながらコミュニケーション上は相互行為性を前提としないし、実際になされているやりとりの多くは相互行為の従来定義を満たさない⁵⁰。ラングトンによる分類を当てはめるならば、対面状況におけるターゲットとなる集団への攻撃性を持つヘイトスピーチによる直接の害というよりは、むしろプロパガンダ的ヘイトスピーチの流通・蔓延による（主には非ターゲット集団における）集団形成及び同集団内の

⁴⁷ 横浜地裁川崎市部決平成 28 年 6 月 2 日など、リアル空間における本邦の事例群の多くで見られるヘイトデモ型については、個人の人格権の一部に含まれる平穏に生活する権利、自由に活動する権利、名誉信用を保有する権利等、広範な利益を民事上の損害として認め、差止めを認めている。他方で上記デモがなされた川崎市でその後に制定された川崎市差別のないまち人権条例（2019 年制定）ではヘイトデモについての罰則（第 12 条）が付されているが、あくまでも公共の場において面前でなされるものに限定され、罰則に至るまで慎重な手続きも規定している。このような限定に現れているように、刑事法領域において、一般的に平穏な生活環境を保持する利益や自由に活動する利益が（刑事的制裁を背景として実現されるべき価値として）直接の保護対象とされるかについては疑わしいものとする。

⁴⁸ 例えば、2018 年、オンライン上にブログ記事を投稿し、在日コリアンの少年に対する誹謗中傷が侮蔑罪に当たるとされ、科料 9000 円の略式命令に至った事件や、脅迫の嫌疑での書類送検後、神奈川県迷惑防止条例違反（つきまとい行為禁止）で罰金 30 万円の略式命令が出されたものなどが報道されている。

⁴⁹ 典型として、京都朝鮮学校公園占用抗議事件（最終 2012 年 2 月 23 日）がある。

⁵⁰ なお先述したゴフマンが述べている例のように、どちらかといえば、自室のテレビ画面の前で悪態をつく、フレームの誤作動状況に近いかもしれない。この悪態をついているものの発話が意図的に発話者に届く場合を別として、偶発的に漏れたとしても、それ自体は、流通の問題である。

差別・偏見の助長を通じた間接的な害に属するだろう⁵¹。オンライン上のヘイトスピーチの多くについては、従来のような集団形成が仮想的にのみなされ、他方ではそうした集団形成を通じた害の契機もまたそれへの抵抗の契機も、コミュニケーションの媒介者やアーキテクチャの介在とともにさまざまになされるためである。こうした害の特性と抵抗の支援手法については論考[展開]で詳述するものの、この相互の**コミュニケーション形式の特殊性**については、リアル空間での議論とオンライン空間での議論を分かつものとして、第一に留意する必要があるだろう。

(2) メディア媒介的なオンライン空間の空間特性

第二に、コミュニケーション形式のみならず、メディア媒介的な**オンライン空間の空間特性**にも着目できる。まず、オンライン空間の特性として、空間的な分割の困難性がある。これはオンライン空間においては全てが可視化されている（あるいは可視化されるべきである）が故にリアル空間を前提にした規制の拡張や規律の一様性を要求するという事態に直結するわけではなく、むしろ適切な空間的限定の下では正当であったはずの要求や法的介入が、その正当な範囲・権限を超えて全域に及びうるという空間特性として把握される。

この典型事例として、鳥取県青少年条例による「指定」を介したネット書店からの書籍締め出し事例がある。具体的には、危険な吹き矢・レーザー等の制作方法を内容に含む「アリエナイ工作シリーズ」等の書籍が青少年健全育成条例上の有害図書へ指定された後、結果としてオンライン書店（Amazon）で販売ページからの取下げがなされ、鳥取県外のユーザーに対しても事実上購入が困難となった事例である。（他でも検索可能な危険物制作方法を書籍に載せることと現実の危害との関係を掘り下げること、あるいは危険情報規制とポルノグラフィ規制との比較を行うことも有益であろうが）本稿の関心からする主たる問題は、この取下げがオンライン書店の自主的な判断と評価するには困難な事情がある点にこそある。すなわち、鳥取県条例が「インターネットの利用その他の方法により鳥取県内において」図書の販売を業とする者が有害図書販売行為を行った場合には罰金・懲役を課す旨を規定しており、オンライン書店としては、購入者の属性に応じて適用される（本邦だけで見ても最大 2000 弱の）地方自治体の条例の規定ごとに販売対象図書をカスタマイズする対応コストをかけるか、書籍自体をページから取り下げる（あるいはアダルト指定する）かの二択を迫られていたはずである。そもそも前者の実施にあたっては、オンライン書店側で購入者の正確な属性を確認・蓄積することが前提となる以上、事業の実施のために必要となる以上の情報収集をオンライン書店側に求める点で背理である。この点に鑑みれば、強いられた選択肢から選択している事実を以て、オンライン書店の自律的判断であるとみなすことには困難

⁵¹ Langton(2018), “The Authority of Hate Speech”の分類が参考になる。なお、本邦においても重なる分類として提案されているものとして直接的・間接的なヘイトスピーチを分けるものがある。那須祐治 (2019)『ヘイトスピーチの比較研究』487-493 頁を参照せよ。

が伴うだろう⁵²。結果として、鳥取県の地理的範囲を超えて、書籍へのアクセス遮断は当該オンライン書店日本語サイト利用者全体に及ぶこととなった⁵³。一言でまとめるならば、オンライン空間においては従来考えられてきた分権性が機能不全に陥った事例として総括できよう。

翻って、オンラインヘイトスピーチ規制についても、同様の地理的な拡張作用が付随することがわかる。オンライン上のヘイトスピーチによる害が個別化されている場合、すなわち侮蔑・中傷や威圧、脅迫、プライバシーや名誉を脅かす虚偽の事実の流布、暴力の扇動等の行為群にあたる場合はリアル空間における規制の処罰範囲拡張や遮断措置も条件的に可能になりうる⁵⁴一方で、害が個別化されておらず、あるいは害が地理的（歴史的・文化的）条件と発話の具体的文脈に服する以上は、発話上、何らかの差別的な言辞を含んでいる発話がなされたとしても、（公人に対する批判や権力と結びつく「マジョリティ」属性への批判を想起すれば自ずと明らかなように）それだけでは規制されるべき対象の範囲確定はできず、規制の根拠を欠くこととなる。

さらに、仮に害の発生が、地理的（歴史的・文化的）条件と発話の具体的状況に即して認められたとしても、規制実施を正統に行いうる地理的範囲が規制主体の地理的管轄に依存しているために、それを超えた執行は権限を超えたものとなるだろう。実際、例えば、川崎市における「川崎市差別のない人権尊重のまちづくり条例」では、インターネット上のコンテンツに対するインターネット拡散防止措置（プロバイダ等への削除要請）が定められているが、そこでは該当する投稿が川崎市内でなされたものあるいは川崎市民を対象とする場合などの限定が付されており、また投稿に対しては禁止・罰則は付されておらず、審議会の諮問を経たのちのプロバイダへの削除依頼手続き及び公表措置のみが規定されているにとどまるのは、上記の権限との調整の結果である⁵⁵。このようにオンライン空間の分割困難性（全域性）は、リアル空間を前提にした規制の拡張や規律の一様性を要求する条件ではなく、また分権性の機能不全ゆえに個別条例的介入も極めて限定的な条件の下でしか正当化の根拠とならないのである。

⁵² オノラ・オニール（2016年）『正義の境界』（みすず書房、神島裕子 訳）第5章でいうところの、重要な利益とパーターにされた選択の自律性に関する「（あなたが）拒否できない申し出」の問題である。

⁵³ 無論、こうした人権制約局面における条例制定との関連では、そもそも青少年保護というトピックが（特に法律の留保からの検討を含め）こうした分権性に適った対象であるのかについては、改めて検討する余地はあるだろう。もしもこの検討によって、当該トピックが本来的に国家に属するべき事項であった（にもかかわらず地方にいわば「投げられて」きた）ことが判明するならば、然るべき審議手続きを踏まえた国家の判断を要求することになろうし、反対に（本文に述べたような全域性に由来する影響関係の拡大から）当該トピックが国家にも地方にも属するべき事項でないことが判明するならば、そもそも法的介入を控えるべきことに帰結しよう。この点について、大屋雄裕による講演（2022）「“危険”な科学・技術に関する情報の統制と自由を考える」（2022年9月24日）が参考になった。

⁵⁴ 典型としての青少年保護のためのサイトブロッキングに加え、近時話題となった著作権保護のための「漫画村」サイトブロッキング、あるいは侮蔑罪の2022年の厳罰化に、空間特性に応じた介入の拡張とその条件設定のバリエーションを見ることができる。

⁵⁵ 川崎市「川崎市差別のない人権尊重のまちづくり条例」FAQ ページを参照せよ。

<https://www.city.kawasaki.jp/templates/faq/250/0000128378.html>

(3) 空間特性を踏まえた「不快原理」の位置付け

以上の分割困難性(全域性)という特性に加え、オンライン空間の更なる特性として、検索行為やフォロー行為を介して(原理的には可視化・開放されつつも)主として自らの意思の下において、さらにはアーキテクチャ上の制約・時空的距離においていわば「閉じられている」点が挙げられよう。

前者の意思性については、オンライン空間の特性上、自らにとって「害と感じられるもの」に接触する起点も(あるいは回避しなかったという事後対処においても)自らの意思によるところが大きい点が指摘できる。この意思・状況依存的な発生過程を無視し、等しく「害」として取り扱うことは、不快原理(offence principle)が陥るのと同じく過剰規制問題を生じさせることになるのである。そもそも不快原理の提唱者であるジョエル・ファインバーグが同原理の適用にあたり、不快の中でも、程度として深刻かつ一般的なものに限定して問題となる不快を捉え、その大規模さや継続性に加えて、道理にかなった回避可能性(容易性)及びその不快さを同意の下で引き受けたか否かを考慮に入れていたことが、ここで想起されねばならない⁵⁶。それと知りつつ自らの意思で近づいた不快、すなわち引き受けられた不快を危害類似のものとして構成するのは、危害に属する範囲を著しく拡大することになるためである。

いわゆる「囚われの聴衆」に関する拒絶困難性を構成する事情においては次の事情、すなわち文字情報はもとより画像情報でさえ強制的に耳に入ってくる音声等と比して態様としては一瞥の後に破棄することが可能であることや、その発信者が公的地位を代表するものであること(政府言論性)、または物理的閉域(あるいは移動の目的のために相当時間の物理的拘束を強いられる公共交通機関車内や生活の糧を得るための職場などそれに類した空間)における強制的・反復的表示機能により避けることができないことといった事情が考慮されることを、ここで思い起こすことができるかもしれない⁵⁷。確かに検索や SNS を通じて「目に入る」という点においては非意思的な側面が強い(いわばテレビ的である)と言えなくもない一方で、文字・音声・動画を含めてコンテンツと向き合うプロセスにおいては回避や統御が可能な、意思的な側面が強く(いわば新聞・雑誌的である)⁵⁸、翻って非意思的

⁵⁶ Joel Feinberg(1984), *Offense to Others (The Moral Limits of the Criminal Law)* の第 8 章[Mediating the offence principle]を参照せよ。

⁵⁷ 横大道聡(2013年)『現代国家における表現の自由—言論市場への国家の積極的関与とその憲法的統制』第 11 章[政府言論の憲法的統制]及びその前後の関連する章を参照せよ。「囚われの聴衆」法理の適用場面は、私人の「表現からの自由」を守るための政府の規制(利益)の正当化場面である。そうである以上、「囚われの聴衆」法理は、例えばその発信者の地位(政府言論性)、場所(物理的閉域性)、態様(非文書=音声性)等により自ずと限定される。自ら検索して、あるいは自ら構築した TL を通じて特定の発話(文字情報)が視界に入ることは、あなたと Google との、あるいはあなたと Twitter との共同制作物である。この結果として、オンライン空間において情報の高度の流通性が仮に存在するとしても、ジェレミー・ウォルドロンという「社会的環境の汚染」、すなわち尊厳あるいは包摂性という公共財に対する計算された攻撃の量的拡大とみる見解は、表現のオープンさを性急にパブリックな影響関係へと「翻訳」している点で妥当ではない。

⁵⁸ 名誉毀損における「社会的評価低下」判断にあたっては、紙の上に固定化・活字化された新聞・雑誌等の文章メディアについては「一般読者の普通の注意と読み方」基準が、音声・映像情報が次々と流れていくテレビ放送については「内容・印象等の総合考慮」基準が適用されてきたことが、ここで参考になる。特に後者については判例上、「録画等

な遭遇においても続いて意思的なブロックが可能であり、被ブロック状況においても意思的に対象者を見ることさえも支援されている。

続いて後者のアーキテクチャ上の制約・時空的距離については、まず同空間が必ずしも参加者全員が共有する「一つの」同じ空間ではない点が指摘されねばならない。すなわち、参加者はオンライン空間において別々の社会的現実あるいは「世界」のヴァージョン⁵⁹を制作しながらも、しかしそれぞれが同じ世界を共有しているかの想定を構築する側面を持つという擬似同期性が指摘できるのである⁶⁰。すなわち、一見すると、トレンドを通じた関心のまとめあげや「炎上」事案は、対面的同期性あるいは「純粹同期性」⁶¹の拡張のように感得されるわけだが、この経験的側面はむしろ上記の擬似同期性を覆い隠すアーキテクチャを通じた選択の結果であり、対面状況とは隔絶した性質を持つ、という点である。(トランザクションは別として) 相互行為の前提となる相互認識は欠如しており、こちらが見ていることはあちらには見られておらず、逆もまた然りである。すなわち、単純に想定されるような無媒介の相互行為や受動的な情報接触とは性質を異にするという意味で、想像的には集合的でありながらも実践的には非相互的な「制作」という側面を持つのである⁶²。

2-2. 害の発生機序を明らかにすることと可罰性との間の距離

(1) ヘイトスピーチ規制の根拠として提示される「害」と「安心」

さてこれらの差異は、本稿の主題であるオンライン上のヘイトスピーチにはどのように当てはまるだろうか。あなたがターゲットとなるマイノリティ集団の一人であると仮定した場合、リアル空間では自らの住まう概して平穏な生活上の世界に突如としてヘイトスピーチに属する発話群が畏怖を与える形で介入してくる(無論、リアル空間においてもよりマイルドな介入として「拒否型」があるが、これはオンラインでは起こりづらいかもしれない

の特別の方法を講じない限り、提供された情報の意味内容を十分に検討したり、再確認したりすることができない」特性が考慮されていたことは、オンライン上の表現の解釈に当たっても参考になる。以上につき、村田健介「インターネット上の名誉毀損と民法法理」法学セミナー66巻12号(2021)を参照せよ。

⁵⁹ ネルソン・グッドマン(2008)『世界制作の方法』(ちくま学芸文庫、菅野 盾樹・中村 雅之 訳)

⁶⁰ この点につき、動画共有サイト・ニコニコ動画のコメント機能に即した擬似同期性の分析として、濱野智史『アーキテクチャの生態系』及び村上裕一『ゴーストの条件』第二部第四章・第六章の分析から示唆を得られるかもしれない。ヘイトスピーチの蔓延状況を網羅的な機械的走査から把握しようとする試みは、社会的な言説を取り出す点で有益ではあるが、少なくともここで問題となっている各人の認識を介して生じる害を明らかにすることは異なっている。ここでの関心は、世界の中にある言説一般の悪性を除去することではなく、社会的状況の支援を受けつつ、オンライン空間の特性の中、個別的な認識の上で現出する害への規制の発動条件に向けられている。グッドマンの言葉を借りれば、既成の世界との対応としての真理ではなく、バージョンの制作による世界制作とその適合性評価として捉えることができようか。

⁶¹ 同上。

⁶² もちろん、ここにはグラデーションがある。一方の極には、Zoom 面接などが典型であろうがリアル空間における対面コミュニケーションと寸分違わぬコミュニケーションを(技術的にも、あるいは規範的にも)実現するものがあり、他方の極にはリアル空間とは異なる空間・時間・関係フレームの上で全く別のコミュニケーションを実現するものがある。中間にはLINE等の既読機能などがあるかもしれない。

63)。その一方で、従来のマスメディアやオピニオンリーダー等を介してなされるのではない、オンライン空間における多くの発話群では事情は異なりうる。まず、この介入は自らの意思が何らかの起点を成している社会的現実あるいは世界制作の結果として現出することとなる（害の意思的選択への依存性）。そして先に「メディア」媒介性と呼んだ効果の帰結として、対面状況では突如としておそってくる、畏怖を伴う、あるいは無用な議論に巻き込まれる等々の性質は、主にはその非直接的な物理的距離な距離を介して失われることとなえう（害の間接性）。またそもそも発話とその受領とが、疑似同期的な錯覚を与えつつも、原理的には隔たっているという時間的距離を基礎として、各種遮断の経路・手段を複数選択しうる人が多い（害の回避手段の選択可能性）⁶⁴。これを支える事情として、反対にリアル空間での署名（集団形成）やデモ（あるいは対抗するカウンターデモ）、つきまといや粘着行為（あるいはこれらの回避）には多大なコストがつきものであったのに対して、オンライン空間での同様の行為への対抗措置は（もちろん、悪しき例として掛け金を支払わない事による過剰さに辿り着きうることはある種のキャンセルカルチャーの例においてみられるわけだが⁶⁵）比較的に小さなコストで済む点も挙げられよう（害への対抗手段の選択可能性）。

以上のコミュニケーション形式の特殊性及び空間特性から、オンライン上におけるヘイトスピーチの多くは、「具体的に特定可能なターゲットに対して、直接的な害を与えたり、あるいは個別的な害を与えたりするわけではない場合におけるヘイトスピーチ」⁶⁶に属するとする見解が、法規制の観点からヘイトスピーチの害を論じてきた多くの論者に共有されている。少なくとも個別の発話単位でオンラインヘイトスピーチを見た場合、その「害」に基づく自由制約の正当化（危害原理）が機能不全に陥るのは、上記のコミュニケーション形式の特殊性及び空間特性の多元性・多層性を規制文脈に置くための「翻訳」に失敗してきたためである、といえよう。より一般的に、法学的なヘイトスピーチ規制（消極）論が、たとえば社会学的な規制（積極）論と少なからぬ緊張関係に立ってきたのも同様の理由に基づく

⁶³ 若林美奈（2021）「集团的差別的言動と不法行為」は、ゴルフクラブ入会拒否や賃貸物件入居拒否など不平等な取扱いをなす類型を拒否型事例と整理した上で、私的自治が一定の範囲で制約され、その限度を超える場合の効果として契約無効及び人格権侵害による不法行為を導く。さらにこの拒否型と区分される類型として、平穏を突如として乱す（典型的には本人に対して名指しで行われる）排斥・危害告知・誹謗中傷をなす差別的言動型のヘイトスピーチの性質を取り上げている。この区分自体は妥当であろうが、他方で事実上の公開性と事実上の匿名性（という想定）に支えられているオンライン上においてはそもそも属性開示が不透明である以上は拒否型は起こりづらく、また平穏を乱す性質も対面状況とは異なることは念頭に置いてよいだろう。

⁶⁴ この性質から、オンライン空間をリアル空間における公共空間と同視したり、あるいはそれを前提としてパブリックフォーラムとしての制約除去根拠として用いる見解は、空間特性の把握において問題を持つ。関連する点を指摘するものとして、土井翼（2023）「SNSとパブリック・フォーラム論—非物理的パブリック・フォーラム論の無用性」情報通信法学研究会 メディア法分科会（令和4年度第2回）報告資料を参照せよ。なお関連して、土井翼（2022）「傍聴人による議事の撮影を許可制とする県議会規則及び映画製作会社による撮影許可申請につき個別事情を考慮せずにした不許可処分が合憲かつ適法とされた事例〔那覇地判令和2年8月5日判批〕」自治研究98号7号）も上記観点から示唆に富む。

⁶⁵ 大屋雄裕（2022）「キャンセル・カルチャーが孕む二つの文脈」Voice（537）148-154の[賭け金は何か?]の項を参照せよ。

⁶⁶ Robert M. Simpson “Is Hate Speech Punishable?”（2023）「ヘイトスピーチは罰しうるか？」永石尚也・本田康作訳、『差別・ヘイトスピーチ・法』（近刊）第2章

ものである。

実際、すぐ上で引用した R. M. シンプソンが自身の論考中⁶⁷で詳細に検討していたように、害の原因に着目した規制アプローチ---第一に、ヘイトスピーチの「行為」を煽動同様の verbal conduct とみなすアプローチ（シンプソン論考の第3節）、第二に、ヘイトスピーチの害の「程度」は、究極的には「感情」への害のため小さいとするアプローチ（同第4節）---は、リアル空間のヘイトスピーチを超えてオンラインヘイトスピーチに対して適用するにあたっては、上記検討の結果として、なおのこと見込み薄である。

これに対して、言説空間における流通・蔓延を介した尊厳への害に先駆的に着目していたとするウォルドロン⁶⁸の議論---ヘイトスピーチの害は構造的に社会において対等な地位を有するという尊厳が保障されているという正当化された自尊心を奪うとするアプローチ（同第5節）---は、オンラインヘイトスピーチに対しても一見したところ重要な示唆を持つように思われるかもしれない。すなわち、ヘイトスピーチが、特定集団に属する対象者に認識させることがすなわち社会の中で平等な地位（尊厳）を持つ「安心」を奪うことだとする議論である。

しかし、この種の議論⁶⁸に対しては、ヘイトスピーチの態様やその文脈（特に差別的慣行が制度的裏付けを有しているのか、あるいはその慣行に対抗する是正作用が制度化されているのか）によって、当然毀損される安心は質的・量的に異なるはずであるにもかかわらず、その存在自体が安心を奪うとするのは、実証的な社会科学的知見を避ける（side-stepping）ものであるとする、シンプソンからの批判があった（同第6節）。コミュニケーション携帯や空間特性上、多元的・多層的な構造を持つオンライン空間では、発話による尊厳毀損の態様やその文脈もまた変容を被らざるをえない。この差異を分別し、各々について実証的に明らかにすることなしには、いわば観念的にオンライン上のある発話をリアル空間上の同内容の発話と、とりわけ万人に開かれた公共の場における同内容の発話と同視する誤りに陥るだろう。すなわち、主たる問題は同視するための条件であり、その同視し得ないケースをグラデーションをもって分別する作業である。このことも、上記の消極論を支えてきた背景として想起されるだろう⁶⁹。

⁶⁷ 同上。なお、同稿は脚注1)で示した2020年初頭における報告原稿として提示されたものに、若干の修正を施したものである。

⁶⁸ Catharine A. MacKinnon(1996), Only words などが代表的であるが、表現の自由に重きを置く論者の中でも、ナイジェル・ウォーバートン(2015)『「表現の自由」入門』(岩波書店、森村進・森村たまき訳)でも、「ポルノグラフィ」表現について、その暴力との結びつきを理由として、実証的な検証を待たずに規制する可能性が示唆されている。

⁶⁹ なお、この批判は、個別のヘイトスピーチそのものを原因とした害の発生に直接着目するのではなく、社会における既存の不平等が維持・継続される状態へと当該発話がどれだけ寄与しているか、その徴候(symptom)としてヘイトスピーチを取り扱う提案に根ざしたものである(第7節)。ウィトゲンシュタインに由来する「徴候」は、論理的な必要条件を示すものでも十分条件を示すものでもないものの、因果に関する経験命題の「確証」を促しはする。ヘイトスピーチが害を作り出す(単称的)因果関係は複雑かつ多重的である一方で、社会における既存の不平等のメカニズムは社会科学的分析に適していることに照らしても、方法論的に見てシンプソンの提案それ自体は一貫したものである。ここから、経験科学的な研究によって明らかにされるだろう、実際の言語の働きや私たちの持つ価値の上げ下げの実態にシンプソンが期待を寄せ、ウォルドロンを批判するのは自然である。

(2) 害の発生機序を明らかにすることと可罰性との間の距離

さて、前述した経験科学的な研究との協働を目指すことは、オンラインヘイトスピーチのもたらす害の解明においても必要であることに疑いはない。しかし、そもその問いであり、シンプソンもまた自身の論文タイトルにあげていた Is Hate Speech Punishable? という (刑) 罰の執行を念頭に置いた規制文脈に関わる問いに戻れば、ウォルドロンのみならずシンプソンその人における回避 (side-stepping) にも着目せざるを得ない。その回避とは、ヘイトスピーチの「可罰性」に関する問いが宙吊りにされていることである。

これはシンプソンのみならず、より広範な論者にも同様に当てはまる。経験科学的な研究との協働を実現するにあたり、この「可罰性」概念を用いることでいかにヘイトスピーチとして (刑) 法によって名指される対象が限定されるのか、その反射的結果として「可罰」的ではないヘイトスピーチに対してはいかなる法的介入がなされうるのか、ひいては可罰にかわるより適切な法の効力 (執行) とは何か。これらについての分析は、ヘイトスピーチ固有の害を解明するに留まらず、(刑) 法的な「可罰性」を問う以上は、避けて通ることができないためである。

それにもかかわらず、シンプソンが法の効力について述べた箇所は、(個別的で直接的なヘイトスピーチについては法による (刑) 罰とその執行を自明視しているのと対照的に) 極めて限定的かつ抑制的である。すなわち、「法には表出効果がある。法によるヘイトスピーチの禁止とは、ヘイトスピーチの根底にある態度や考え方に反対の立場をとることを国家が伝達する効果的方法となりえるのだ」とするにとどまるのである (同第 6 節末尾)。

これはいわゆる法の表出効果に着目するものであるが、表出効果として典型的に論じられるネガティブな側面である萎縮効果や適正手続保障なしの同調圧力に明示的に触れることない点に、法理論の観点からは明示的な不足がある。すなわち、オンラインヘイトスピーチの上記特性に即して規制の可否を論じるためには、従来型の個別の発話による望ましき／害の正確な把握とともに、その集積としての言説空間における流通・蔓延の望ましき／害を、その社会的文脈自体の特性として実証的に明らかにした上で、規制の文脈に接続する必要があるのである。

オンラインヘイトスピーチを含む特定の種類の発話の可罰性 (Is Hate Speech Punishable?) を問うということは、単に発話の害を一般的に問うことではなく、他ならぬ (刑) 法で罰を与える条件を問うことである。すなわち、可罰性なる概念を具体的事例に適用する営みは、日常的な意味での道徳的非難によって追認されるものではなく、当該事例を取り巻く具体的な (刑) 法のコンテクストとそれを持つ社会のコンテクストを明らかにする作業を通じ、執行の過剰と過小をともに排除することで初めて正当化される。

言い換えれば、可罰性を問うことは、少なくとも当該言論に由来する法益侵害を特定する (従来考えられてきたそれが狭きに失するのであれば、概念的に拡張する) ことであり、当該言論が許容 (免責) される条件を整備することであり、当該言論を適正な司法手続に沿っ

て裁定に導く制度を検討することにほかならない。これらを明らかにする作業なくしては、（およそ完全に不平等が解消された、完成された社会なるものが観念しがたい以上）いみじくもシンプソンが主張する経験科学的研究が提供するエビデンスの匙加減⁷⁰で、一般化されたヘイトスピーチ対策としての禁止と、その結果としての「均質化された世界共同体」（同第1節）と呼ぶ状況への傾斜を、期せずして果たしてしまうためである。

冒頭で確認したように、機能的に把握されるほかない「ヘイトスピーチ」自体が強い文脈依存性を持ち、一般的に定義することの困難を抱えている。これはヘイトスピーチの概念にまつわる哲学的な困難である。しかし、法（哲）学的にはさらに、当該文脈依存性を持つ法的判断の適切な利用と濫用、判断する主体の重畳性、当該判断における適正手続等の法的介入への統制の問題を避けることはできない。ここから、経験科学との協働を期するためには、シンプソンらの回避を埋める法（哲）学の仕事として、この「可罰」性概念を取り巻く、罰、主体、手続にかかる諸概念とその統制原理を、先立って明らかにする作業が求められる。続く3では、この法的介入及び不介入への統制原理を（ごく図式的なものに留まるもの）取り出すこととしよう。

⁷⁰ 林岳彦（2023）「Evidence Based Practice にとって「良いエビデンス」とは何か」に倣い、エビデンスを文脈特定性や社会的妥当性を超えて過剰利用するケースを念頭に置いてみよう。エビデンスそのものの科学性と政治性が両立しようように、エビデンスの科学性とエビデンス使用の政治性とは両立するため、経験科学的研究が参照される場合にはその利用の適正性が鋭く問われる。井頭昌彦『質的研究アプローチの再検討』（勁草書房）所収の同論考を参照せよ。

3、ヘイトスピーチの害と規制による害との各種リスクトレードオフ

3-1. リスクトレードオフ①：法的介入によるリスクと、不介入によるリスクのトレードオフ

(1) 法的介入を統制する「法の正当性」

繰り返すように、ヘイトスピーチのある種の形態は、個別的な権利あるいは集団的な「公共善」の観点から特有の害を持つ。このことは疑いないし、その害をもたらすメカニズムを、言語が持つ（発話的 locutionary, 発話内的 illocutionary, 発話媒介的 perlocutionary な）発話上の諸特性や社会の具体的状況から導くことも有益であるだろう。

しかし、特定の発話類型が構成／発生させる害の除去手段として法による規制が候補に上がる場合には、後者についても少なくとも同程度の相等的な注意が要される。端的に言えば、**法それ自体が言語的側面から（発話的・）発話内的・発話媒介的な害をも有するため**である。強盗を働くガンマンが「金を渡せ、さもなくば撃つぞ」と命令するのとまずは同様に、制裁付加を伴ういくつかの法もまた、その言語上の効力を用いて、害悪の予期を背景とした命令に強制力を付与している。

とりわけここで焦点が当てられている（刑）罰の執行を基礎付ける刑法の犯罪規定は、まづもって法に関連する公的な職務従事者 officials を縛るルールでありつつ、特定の行為を個別の行為者に命じる発話内的な illocutionary 指導性をも併有するとともに、解釈権限の非対称性を背景として、制裁としての（刑）罰すなわち物理的害悪の潜在的発動を告知するものである。さらに、ある行為が（刑）罰の執行と関連することを示す法文は、その発話媒介的な perlocutionary 効果として、それが少なくともある時点におけるある社会的・歴史的背景の下での共同体の決定であることを承認するよう迫り、またそれ自体が当該共同体の基盤的規範に他ならないとの錯覚を与える効果をも持つなど、多岐化したさまざまな効果を伴うだろう。

この結果として、およそ何らかの害悪に対抗する法は、法自身が作り出す害を最小化する仕組みを備えるべきことになる。すなわち、ガンマンの命令と法の命令は次の点で明確に異なる⁷¹。そもそも法の存立において、制裁付加を伴ういくつかの法は、その命令が従われるべき責務を伴うことになる⁷²。そうである以上、法は自身が正統な権威を持つことを主張する必要に迫られる⁷³。単なる暴力行使者であるガンマンにはそのような主張は必要ないわけだが、刑罰の執行は特有の条件の下でのみ発動する、正当化された暴力行使であることが必

⁷¹ H. L. A. Hart が明示的にこの点を明らかにしている。『法の概念』第3章を参照せよ。

⁷² Andrei Marmor(2010), *Philosophy of Law*, ch1 を参照せよ。本文の事態は法の構成的ルール性に由来する。ガンマンがその発話において達成する単純な財の移転とは異なり、法はチェスなどと同様に、ゲーム内における制度的な役割を構成するとともに、ゲーム内において制度的な役割を果たす play 人々が現に従うルールを構成するルールによって統御されているのである。

⁷³ Joseph Raz, “Authority, Law, and Morality.”を参照せよ。もちろん、そうではない法として技術的な法分野についての安藤馨による指摘があるものの、この点については前掲注)20 永石「法の執行と正統性」第4節を参照せよ。

要なのである。より噛み砕いて言えば、制裁の告知そのものが国家による公布と運用という（主体及びその作用において）特有の発話内 illocutionary 行為を介した潜在的な物理的暴力行使の予告であり権利制限であることが避けられない以上、その暴力行使・権利制限がその形式において理解可能・遵守可能であり、さらにその内容において正当化されたものであることがともに要求されるのである⁷⁴。

例えばこの内容上の正当化にあたって典型的には上位の（憲法）規範が呼び出されることとなり、可罰性の概念は、少なくとも「どういう場合に刑を科して良いか」、「どの程度の刑を科して良いか」、「どのような手続によるべきか」、「刑を科す場合や刑の重さをどう定めるか」などにかんする限定・条件とともに問われることとなるだろう。オンライン上の表現に即して言えば、上記の法文を介した事後規制に限られず事前規制として、特定の表現については（限定された）事前規制・事後規制を通じて発話 locutionary それ自体の削除・不可視化措置すらもアーキテクチャの支援を介して広範に可能となる以上、比例的にその内在的限定が求められることとなるのである。

このように、ガンマンの発話の害とは異なる、法の權威を支える法の害の除去原理には、第一に、法的介入を統制する「法の正当性」要求がある。

（２） 法的介入を統制する「法の正統性」

さて、上記の意味で（刑）罰の執行の正当性が認められた場合にも、他ではなく（刑）罰執行を必要とするには、まだ間隙がある。權威を伴う法的制裁の発動にあたっては、技術環境や社会環境の変化の中で、むしろその効率的実現やその機械的实现によって、尊厳に反する選別、選択的執行や panvasive な可視化に由来する自由の領域の縮減など、個人に必ずしも帰属しない別の水準の害をも生じさせることが知られている⁷⁵。すなわち法は、権利の語法に解消されないその害を限定するためにこそ、制裁の発動そのものを最小化するための仕組みを整備してきた側面もある⁷⁶。制裁の極限たる（刑）罰における限定原理が、憲法上も独自の位置を持つとともに、技術的・社会的環境の変化に伴い「翻訳」を要するのはこのためである。

すなわち、法的權威を基礎付ける第二の点として法的介入を統制する「正統性」要求と呼ぶべきものがある。当該の害への対応がそもそも法に求められるべきか（法とその他領域との区分）、仮に求められるとしてどの法領域における責任・制裁システムと結びつけられるべきか（民事・刑事の別）、害の量と禁止が単純に比例するべきか（国家の作用と権力抑制

⁷⁴ 憲法第 31 条から第 40 条がメタ的な統制原理の端的な例であるが、無論原理は明文化されたこれら条項に尽きるものではなく、法の支配や法内在道徳などの法理念からも導き出すことができよう。

⁷⁵ いわゆる「法を破る自由」の議論やジュリーナリフィケーション、あるいは「ハックする自由」などを参照することができよう。また panvasiveness については前掲注)6 永石「公共空間における情報識別と警察活動についての覚え書き」を参照せよ。

⁷⁶ 例えば、個人データを用いた自動処理として個人を選別して取り扱うことは、ドイツ行政手続法上など禁止される例が散見される。より一般的に、行政活動における責任主体を名指しするために、自動執行や匿名的実行の排除や理由付与手続きが充実してきたことなどにも、これを見ることができる。

の均衡問題)、この実現をどの部門に求めることが正統性を持つか(司法・立法・行政の抑制均衡)、その帰結としての制裁の過剰を部門ごとに回避する枠組みとは何か、制裁の過小へに対応する部門間のコミュニケーションはいかに組成されるべきか……これらの課題に応えることこそ、具体的なヘイトスピーチへの(刑)罰の執行が正統性を持つための条件である。

(3) 法的介入の統制と不介入の統制

さて、このように正当性・正統性による統制がなされるのは、法的介入のみならず、法的不介入にも及ぶだろう。法を通じた権力作用には、介入すべき場面で介入しない不作為、介入場面において手続きを履践しない不作為、さらには介入場面において現実に介入する対象の恣意的選択(選択的執行)もまた含まれる。

本稿の主題である、個人的法益とは隔絶された特定集団へのヘイトスピーチ規制に即して言えば、現実の執行段階で宣言された法が(正当性と正統性の観点から)拡大的にも縮小的にも曲げられ、ターゲットであるマイノリティ会員にとっての負担となる形で運用される危険は、他の社会的不均衡を前提としない法律以上に強く存在することが推定される⁷⁷。こうした種々の不作為の場面では、法を通じた対処によるリスクとともに、法の不対処によるリスクの間を埋めるべく、最低限、宣言された法とその権限行使の現実との一致⁷⁸が要され、その監視の実質化が手段として組み込まれる必要があるだろう。こうした対処と不対処の間のリスクトレードオフの中に、ヘイトスピーチの可罰性にかかる問いは置かれている。

以上で見てきたように、法による特定の害の除去と法による特有の害の生成とは、原理的に同時的であり、一方のリスクを除去することは他方のリスクを生み出すトレードオフ関係にある。さらに、トレードオフ関係は、法による対処と不対処の間でも生じる。これらの結果として、種々の害に関する二階の決断と、決断を統制する法特有の原理としてその正当性と正統性が求められるのである。

3-2. リスクトレードオフ②: 法-外のものの対処によるリスクと、法-外のものを法が統制す

⁷⁷ 論考[展開]第2節・第3節で詳細に比較するが、例えば本邦において、法の効果としての責任に着目するならば、特定人を対象とした、具体的な害を伴う差別的言動については、刑事上は名誉毀損罪・侮辱罪を構成するとともに、民事上は不法行為としての責任を問われる一方で不特定多数の集団に対する差別的言動については、法益となる外部的名誉・主観的感情の不定性と、その認定上の困難・恣意の危険から、本邦においては直接に規制の対象とはされていないし、その民事賠償も困難である(山本敬三(2013)「差別表現・憎悪表現の禁止と民事救済の可能性」国際人権24号を参照せよ)。しかしこの区分は単なる規制の不足を示すものではなく、(法)規制の正当性の問いと正統性の問いに答えた結果でもあるのである。なぜなら、①集団に対する差別的言動については、個別の法益を具体的に侵害するケースとは質的に異なる上、②輪郭が明確かつ限定された個人的法益とは異なり、集団の範囲設定によってはむしろ当該集団にとって不利となる規制をも招くリスクさえあるためである。

⁷⁸ いわゆるフラーによる法内道徳の第八原理、すなわち congruence between the rules as announced and their actual administration (the rules as administrated)は、この意味で他の7つの原理とは異なり、法の執行・実現段階における統制を読み込みうる唯一の原理である。ロン・フラー『法と道徳』219頁を参照せよ。

るリスク

(1) 自主規制を通じた統制回避リスクとヘイトスピーチ解消法

さて、これら正当性の問いと正統性の問いを明示化することは、翻って法を通じてなしうるコントロールの幅を、法以外の別の諸力、すなわち規範、市場、アーキテクチャを筆頭とする権力作用に対する「間接規制」へも広げる点においても意義がある。自主規制等何らかの形で法執行が私的主体に委ねられる場合を考えてみよう。インセンティブ構造上、執行機関には執行段階における法介入への統制を逃れるインセンティブが生じることは避けられない。このため、当該の統制回避リスクを（委任の段階、執行の段階、管理の段階、終局的には裁定の段階において）包摂する補完原理を付与する必要がある。

では、法が非-法的な領域における作用を後押しし、同時に制約することは、どのような条件の下で可能であろうか。この一つの例を、本邦におけるヘイトスピーチ解消法に見ることができる。

同法では、対象となるヘイトスピーチは要件上、限定的に定義された上で（同法第二条）、効果としても可罰的なものとはされておらず、実効性確保手段も定められてはいない。ただし、これは規制の不徹底を意味するわけではなく、上記の法による害を最小化した結果でもあるのである。

この結果、ヘイトスピーチ解消法は、専ら公共空間における相互尊重・互惠性を促進することを企図したメッセージを発する象徴的な表出効果により、法システムにおける重みづけを果たすとともに、人々の行動変容を促すに留まる。反面から見れば、むしろ規制としてのソフトさによって、権力の暴走を抑えると共に社会における公共討論を保障する機能を有するといえよう。さらに個別の攻撃的な表現者を排除するのではなく、その表明に対する受領者としての後ろ盾となる自尊感情と社会的承認の力とを、共に促進する機能をも果たすのである。言い換えれば、思想の自由市場からの防御が手薄になるマイノリティ成員の発話を支援する領域を開き、「より多く」の発話（＝対抗言論）を促進する機能と言い換えうるかもしれない。

他方で、こうした指針提示そのものにも問題がないわけではない。定義されたヘイトスピーチを違法にするわけでも処罰するわけでもないとは、裏返せば「適用を欠いた法」であることを意味する。これは法の支配と対置される「空気による支配」の問題そのものであるし、そうではなく仮に手続的な統制なく準-法的な裏付けを付与するものであるならば、古くはロン・フラーが論じたところの選択的執行（selective enforcement）及び威嚇執行（jawbone enforcement/enforcement by blackmail）という争訟不能性と執行者による介入の恣意拡大

⁷⁹が間接規制段階において再来することとなる。

このようにヘイトスピーチ解消法には、方や権力抑止による市民的連帯の促進という側面と、方や市民的連帯そのものの暴走抑止機構の不在という側面が緊張関係を伴って混淆している。しかしそこは、(主体、内容等において)法規制に服する表現が定められる公的領域でも、他者を排した純然たる私的領域でもなく、その中間にある独自の領域としての意義を持つ。公共圏⁸⁰あるいは緩衝地帯⁸¹とも呼べようこのグラデーションを持った中間領域は、法を形成する基盤でもある多層的な背景からの声を響かせることで、個性の発露としての声の発明とその批判としての議論提起、そして議論を通じた異なる「最終の言葉」を持ちあった他者への寛容とを促す場であり、コンフリクトを継続的に保持できる環境の維持こそが求められる領域である。

(2) 法介入の領域とは異なる表現の中間領域形成とヘイトスピーチ解消法

かつて J. S. ミルが「他人に対する義務の問題」と「自己配慮に属する行為の諸問題」とを劃然と区別した⁸²際、その中間に社会的な支えを要する「状況の多様性⁸³」が必要であると先んじて記載していた趣旨は、ここに求められよう。それ自体としての価値に乏しく、ときに他者からの嫌悪を受けるだろう「醜い自由」⁸⁴は、社会的専制に由来する「魂の奴隷化」⁸⁵を排することにその根拠を持ち、状況の多様性の維持・促進の支えによってその存立の場

⁷⁹ 森悠一郎「ポリティカル・コレクトネスの意義と限界 差別との闘いが孕む差別 (特集 差別問題のいま 法は差別とどう向き合うのか)」(2023) 法学セミナー818号の第4節[差別からの保護における差別]では、ポリティカル・コレクトネスを通じた非社会的非難において、対象の属性によって保護の不均衡が生じる状況を批判的に解説している。そこでは、保護対象の線引きとして歴史的事実としての差別が不均衡を正当化する根拠として取り上げられることがあるものの、歴史的事実は現在の差別的取扱いを特別視させるものではないことに加え、それによる規範的な侮蔑 demeaning の意味付与は決定的とは言えないものとの主張がなされている。さらに別の正当化論拠として、保護対象の線引きは政治的な無力化の程度に相関しているとの主張については、現行の保護対象の線引きが現実のそれらを適切に反映しているとは言えないものと評価されている。

⁸⁰ 前掲注)33 永石尚也「プライバシー・監視・アーキテクチャ」(2020)。ここでは、Th. Nagel と M. Nussbaum の間のプライバシーについての議論を紹介しつつ、公的領域にも私的領域にも属さない、公的でありながら秘匿された「秘密そのものの秘密裏の可能性」の領域を抽出していた。この概念は、ジャック・デリダの『アポリア』に由来する。『アポリア 死す「真理の諸限界」を“で/相”待-期する』(人文書院、2000年)162-163頁を参照せよ。

⁸¹ ハンナ・アレントが『人間の条件』で言及している No-Man's-Land への参照は、本稿での議論に対応するものであると解釈しうる。

⁸² J. S. ミル (1971)『自由論』(岩波文庫、塩尻公明、木村健康 訳)169頁

⁸³ 同上 47頁。これを支える社会的な支えとは、「自己自身が多数者の優位に反抗しつつ、公衆の意見や傾向と相違する意見や傾向を保護することに関心を持つ、しっかりとした社会力」(148-149頁)を意味する。もちろん、ヘイトスピーチ規制の難問は、ミルが「多数者の優位への反抗」と「公衆と相違する意見・傾向」の保護とを素朴に結びつけていたのとは異なり、現象レベルでは主として(例えば日本における日本国籍を持つなど)特定の属性における「多数者の優位」を象徴的に利用することでなされるものでありながら、同時に社会的な支えというメタレベルでは「公衆の意見や傾向と相違する意見や傾向」にも含まれる個性の発露でもありうるという、一見混乱した帰結を生むためである。しかしこの両者は実践的には統合不能なものではなく、メタレベルでの「社会の支え」を通じた多様性を保持する環境の促進を、法がその象徴的な表出効果を通じて間接的に担うことによっても結びつきうるだろう。

⁸⁴ 若松良樹(2020)『醜い自由』(成文堂)5-6頁

⁸⁵ 前掲注)82 ミル (1971)『自由論』15頁

を持つ⁸⁶。すなわち、「自分の嫌悪する（他人の）行為を自分に対する権利の侵害」と取り違え、「自己の感情を傷つける暴行であるかのようにそれに憤る」⁸⁷ならば、その総合であるところの社会的非難は社会的暴虐の遂行となる⁸⁸。法規制を用いてであれ事実上であれこの中間領域を締め出そうとする傾向性が働くことは、社会的専制の必然である。そうである以上、ミルはこれへの対抗として、「差異が存在することそれ自体」の有益性を、その結果の良し悪しやその予測とは無関係に尊重されるべきものとして主張していた。説得や警告はあれどもこの差異を生み出す個々人の個性の発露は、社会的な「状況の多様性」の支えを通じて少数者の声を響かせるだろう。この理念は、ミルの時点において既に形成されていたのである。

もちろん、新型コロナウイルス感染防止における各種「要請」や解釈指針の提示、さらに近時議論されているところでは実効性確保手段としての各種ナッジの問題等、本来は法介入への統制が働くべき要請を、執行側が回避する動きが進んだ。こうした法の支配からの逸脱事例を正統性や正当性から法領域に包摂することは必然である。他方、そもそも法ではなくその市民的基礎を形成するコンフリクトを統御することとは、上記ミルの見解に見られるように全く異なる。非-法的手法が一般的に内包する問題とは、中間領域構成型のヘイトスピーチ解消法が内包する問題は質的に異なるのである。

何よりもまずヘイトスピーチ解消法それ自体は、個別の表現行為・表現の良し悪しの判断へと誘導するものでもない。同法は、個別のヘイトスピーチの「悪さ」を基礎づける公的規範というよりは、（国家ではなく）市民社会のイニシアティブで構成する「空間設計」を志向する公共圏の維持規範であり、同時に、その道徳的過剰を抑止することをも志向する公共圏維持的規範であると言い換えられる。

すなわち、同法は、第一に、同法に反する個別の表現についての判断を事前決定することがないし、同法に明示的に言及する裁判例群も、同法の趣旨に照らして個別法上の不法行為や名誉毀損等の該当性を判断してきたものである。同法における過小包摂があえてなされたものだと考える場合、本邦におけるヘイトスピーチのコアに属するターゲットだけを取り出したことから、本法における対象限定が終局的なものではなく、常に周位的な差別実践を取り込むべき（明示的には誘導することなく）議論に開きうることを示唆していると解釈できよう。この結果でもあるが、第二に罰則に裏付けられた規制を欠いており、専ら市民

⁸⁶ 前掲注)84 若松良樹『醜い自由』153-155頁では、個人の発展可能性としての個人内多様性と、個性への不寛容を拒否する個人間多様性を双方含むものとして整理されている。そして当然ながら、「多様性が価値を有しているとしても、多様性のために今何をすべきかはそれほど明らかではないのであり、政府が直接に多様性を実現することには限界がある」（191頁）にも留意する必要がある。

⁸⁷ 前掲注)82 ミル（1971）『自由論』170頁

⁸⁸ 前掲注)82 ミル（1971）『自由論』15頁。「社会自らが暴君であるときには（社会を構成している個々人の上に集団とひての社会が君臨しているときには）暴虐遂行の手段は、社会がその政治上の公務員の手によって行いうる行為のみには限られていない。（中略）社会は自己の命令を自ら執行することができ、また実際に執行しているのである」。保護されるべき属性をその内容において特定する法規制を求める要請は、反転してマジョリティ側の武器として通用するリスクをも持ち、そのリスクは不均衡な形でマイノリティ側にこそ降り掛かりうる。

社会における攻防における主軸となる価値を表現している（政府言論）と考える場合には、まさしく市民社会においてときどきに変転する「あるべき表現」と「そうでない表現」を競い合わせる空間が維持されているか否か、典型的には言論空間における不平等性・競争排除といったメタ的な「悪さ」が現出しているか否か、という二階の「悪さ」に焦点を当てているものとして、罰則の不在を積極的に理解できる。これら二点を通じた、非規制的で自己抑制性を（ある程度）備えた競争環境維持装置として、同法の限定性は（先に挙げたオーストラリアの1975年連邦人種差別法18c（1995年刑事罰導入）における慎重な手続・運用状況と同様に）積極的に把握できるものである⁸⁹。

ヘイトスピーチ解消法が指針提示するに留める一方で、ヘイトスピーチへの市民的対応を促進する趣旨は、この法以外のものの対処によるリスクと、法以外のものを法が統制するリスクの緊張関係に求められる。法・法以外による対処の間のリスクトレードオフの間に、ヘイトスピーチの可罰性にかかる問いは置かれている⁹⁰。

（3）オンライン空間における表現の中間領域

さて、これらトレードオフ関係を踏まえ、本稿の主題たるオンラインヘイトスピーチへの規制についてのこれまでの議論のまとめとしよう。

発話による諸効果の分析からは、その形態については、コミュニケーション形式の特殊性及び空間特性からの特定と限定を要することに帰結し、それへの規制にあたっては、（刑）罰の執行に向けた一連の手続き自体が生じさせる害を最小化すべき要請が働くものであることがわかった。オンラインヘイトスピーチの形態を厳格に画するとともに、法による執行の過剰と過小をとともに排除する作業が求められるのはこのためであった。つまり、当該発話行為が「可罰的か」あるいは「罰に値するか」を法（哲）学的に問うためには、（刑）罰の執行にあたって前提となる問いを分割する必要がある。

⁸⁹ 翻って、法という特殊な手段で言論空間における競争を上から内容的に改善することは、各種の困難を伴う。前述した権力の暴走抑止はもとよりとして、より基底的には、構造的な不平等・差別による脆弱な発話者・発話内容へのエンパワメントによる改善を抜きに真実性・公共性の判定によってこれをなすことは、却って公共空間における言説空間を目指す共通価値（自尊感情と社会的承認の価値）を毀損することが難点となる。論考[展開]第4節を参照せよ。

⁹⁰ 大阪市ヘイトスピーチ条例公金支出無効確認請求事件（最判令和4年2月15日）は、ヘイトスピーチの対象については明示的に踏み込むことなく、その制限についても、「制限の態様及び程度においても、事後的に市長による拡散防止措置等の対象となるにとどまる」ものであり、「当該要請等に応じないものに対する制裁はなく」、「認識等公表についても……法的強制力を伴う手段は存在しない」とする。この結果、個々の表現行為・目的（+動機）を定義・分類する困難、認定・該当性判断の困難に対して、審査会設置・運用を含むメタ的なプロセスへと課題を移送することを法的に根拠づける点において評価できる。なお、判例評釈として、高瀬保守（2023）「大阪市ヘイトスピーチへの対処に関する条例（平成28年大阪府条例第1号）2条、5条～10条と憲法21条1項」、毛利透（2022）「大阪市ヘイトスピーチ対処条例の合憲性/（最高裁三小法廷令和4年2月15日）」、阿部和文（2022）「[時論] 最高裁第三小法廷令和4年2月15日判決に関する覚書——大阪市ヘイトスピーチへの対処に関する条例の合憲性に関する判断を中心に」ジュリスト2022年7月号（No1573）には初稿執筆後に接した。

公表の効果については、あわせて「平成30年1月17日 大阪市ヘイトスピーチへの対処に関する条例の施行に関する事項について(答申)」7-8頁も参照せよ。

可罰性の根拠についての哲学的探索が、(法)規制の形態としての(刑)罰執行の妥当性を直接には導かない理由はこの点に存する。ある発言が道徳的その他の意味において妥当性を欠くとしても、その発言に対する規制が法的に妥当ではないことは、上記の法的介入を統制する「法の正当性」及び「法の正統性」の観点から十分にありうる。

このことは反対に、(刑)罰の執行について消極的判断に至ったとしても、それ自体としてはヘイトスピーチ及びそれに準ずる差別的発言への放置・容認を意味しないことに帰する。もちろん、論考[展開]第4節で詳述するように、差別的発言への対抗言論を促進する環境なくしては、象徴的には消極的な容認と等価の作用を果たしうることは前提であるが、規律対象の持つ害と規律自体による害との間にある種々の問いは区別して論じられなくてはならない。可罰性の前提にあたる正当性・正統性を左右する考慮事由を、まさにオンライン空間での道徳的意義を持つ生活の変容に即して明らかにすることこそ、ヘイトスピーチの可罰性に基づく(法)規制を論じる手前にある問いなのである。

4、訂正と赦しの空間への法の寄与 キャンセルカルチャーとヘイトスピーチの間

(1) 言論のキャンセル要求とヘイトスピーチとの反転 表現の自由・再考

以上、本稿では、いわゆる表現の自由の保護範囲や規律密度に関する議論とは別のルートで、いわば川下に位置するオンラインヘイトスピーチの発話特性や発話文脈特性から、その禁止や規制を論じる際に解消すべき論点を明示化しよう試みてきた。まとめにかえて、以上の検討順序とは逆向きに、いわば川上に属する規制の正当化可否から不適切・不品位発言へと「キャンセル」作用を及ぼすことについての議論、とりわけ言論の自由市場モデルの議論⁹¹から本稿の辿ってきたルートを再記述し、両ルートの合流する潮目を読むこととする。

さて、国家の権力にも社会的専制にも抗するものとして、他者を害する「おそれ」のある表現であっても、それが「おそれ」とどまる限りは、法制度の中に位置を占めるにあたり、多様性の維持・促進という観点から保護と規制の間の緊張関係の上に立つ。一方には、顕著な害が顕在化していない限りにおいて、相手の口を噤がせるための措置を要求するのではなく、むしろ更なる説得・批判によって対抗する「より多く」の言葉を紡ぎだす営みが推奨される、とする考えがある。他方で、瀰漫的な害の「おそれ」は、ある属性の下でマジョリティではない側に集中しがちであり、対抗言論提示自体が弱化させられうる、との懸念にも相当の理がある。

この緊張関係は、ウォルドロンを紹介した前述の箇所で既に見たとおりである。しかし他方で表現の自由とは、ある時点の、変わりゆく社会の、移ろいゆく文化の限界のうちにおいて「悪い」表現とそうでないものとを切り分け、前者の発言者を「キャンセル」し、あるいは表現そのものに対する「カードを切る」誘惑にかられてしまうまさにその場面においてこそ、切り分けと「キャンセル」誘惑への抵抗として発展してきた。このこともまた忘れられがちである。

先に挙げたミルが、説得や警告が「差異を生み出す個々人の個性の発露」を促す「社会的な支え」の中でこそ意味を持つと主張していたことを、ここで改めて思い起こすことができる。ミルによれば、表現の自由の行使の結果として排除を正当に要求できるのは、私的な「うち(=私)の場でやるな」であるに過ぎず、「(お前の)うちに閉じこまれ」と公共空間からの撤退を強いることではない。なぜなら、その要求は「自分の嫌悪する(他人の)行為を自分に対する権利の侵害」と取り違え、「自己の感情を傷つける暴行であるかのようにそれに憤る」振る舞いの全き反復に陥るためである。発信と受信とは、社会におけるアクセス可能な総量としては一方の増減がもう一方の増減には直結するため、その片方のみを享受することが構造的にできない。言論はこの対称性ゆえに、個人において培われる受忍の限度及び社会において培われる寛容の限界において、絶えず挑戦を受けているともいえるだろう。

⁹¹ 前掲注)41 大屋雄裕 (2020)「赦しと忘却」アステイオン 93号 110-125頁

ウォルドロンは、感情毀損と区別して尊厳毀損を取り出す作業の中で、この問題に答えている。それによれば、尊厳とは、共同体における平等なものとしての地位を指しており、その地位に見合う承認と扱いに対する要求を生み出す源泉である。すなわち、当該地位が持つ客観的で社会的な側面を意味する一方で、恥辱や侮蔑によって傷つけられた感覚・感情は、ヘイトスピーチの害たる尊厳の毀損の徴候に留まる⁹²。無論、その感情がそもそも相互的な尊重に等しく支えられることで成り立つ自尊心などの毀損は社会的地位と感情とをつなぐものであることは疑いない⁹³。しかし、仮に特定のアイデンティティ属性への攻撃を介して生じる不快の感覚・感情をより分けることなく、自己の尊厳への毀損という事態と同一視するならば、「差異と不合意のただ中で社会が運営されることをはるかに困難に」⁹⁴する、とウォルドロンは述べる。必要なのは、客観的に成立する「かのような」社会的な関係を絶えず新たに構築することであり、各々の間の差異と不合意を消去することではない。

アレント及びミルを引きつつ検討した通り、重要となるのは、公私の間にあり、空間的・時間的に厚みを持った、上記の受忍を支え、寛容を育む「社会的支え」であったことが想起されよう。すなわち、個人的な関係解消や説得・批判と全域的な不可視化・沈黙化との間にある、「社会の支え」による抵抗の潜在的可能性を常に保存し、反対に（機会主義的な否認⁹⁵を

⁹² ジェレミー・ウォルドロン『ヘイトスピーチという危害』第5章（126-128頁）を参照せよ、

⁹³ アヴィシャイ・マルガリート『品位ある社会 〈正義の理論〉から〈尊重の物語〉へ』（2017年）第7章を参照せよ。もしも主観的に価値づけているものであれば、アイデンティティ構成的な属性はその自律的選択により、客観的には脱ぎ捨て難くなり、もしも社会的に押し付けられたものであれば、その差別的環境のために主体的に脱ぐことが困難であろう。ウォルドロンが理念的な平等な尊厳によって等閑視するのは、このアイデンティティ構成における主観的であるととも客観的な側面である。アヴィシャイ・マルガリートは、ウォルドロンのいう感覚・感情をさらに区分し、ウォルドロンのいう尊厳へと結びつけている。そこでは社会的名誉を毀損する侮蔑と自尊心を傷つける屈辱とは区別されるが、後者の屈辱は単なる主観的な不快にとどまらない。屈辱は、共同体に属する人間存在を共同体から拒絶すること、すなわち人間存在を普通の人間以下の存在として取り扱う事態に由来するが、そこで自尊心は、自分が傷ついているか否かの主観的感覚にかかわらず傷つく。なぜなら、自尊心の傷つきは、単に孤独のうちではなく、他者が自分自身をどのように取り扱っているかを無視することはできないためである。ウォルドロンを引き継いでいえば、上記感覚・感情の中でもとりわけ自尊心だけは「政治の状況」のうちにあって、それを支える基底としての地位と不可分の共同体感覚を示すのである。

⁹⁴ 同上 159-162頁では、これを「アイデンティティの政治の無責任さ」とし、「ある争点が私のアイデンティティにとって決定的に重要だということによって、私はその争点についての私の見解を政治的に交渉不可能なものとして提示する」ことによる集合的決定の挫折と呼ぶ。

⁹⁵ ここで1-1節に立ち返り、いま一度、蔑称使用・隠語使用による害との比較を行うことは有益であるだろう。蔑称は、発話者が意識せずともそれを使用することによって社会的現実あるいは世界をカテゴリカルに切り取り、ホーリスティックに「制作」する側面とともに、意識的にそれを使用することによって、自らがその語を取り巻く社会的状況においてどの位置にあるのかを周囲に顕示する側面を持つ。もちろんこの二側面は、他の用語（例えば、一見したところ中立的な地理記述的用語である「中東」や「極東」に含まれる権力性など）や慣用句でも一般的にそう機能する可能性はある。その一方で、語に対する訂正過程や語に含まれる脱権力化プロセスを介して、語を自分たちの手元に取り戻す抵抗の動きもまた組成しうるのである。

この蔑称の二側面に対して、隠語はその後者にあたる顕示的側面につき、弁明として「そういうつもりはなかった」という否認の余地を持つ。蔑称については「そういうつもりはなかった」としても特定文脈の下で現にそう使用したことが非難されうるのに対し、隠語はその使用の文脈をより特定状況依存的な形でもち、それゆえに「そういうつもり」の意図が弁明として機能する余地をより広く残す。この点に蔑称と区別される、隠語の特徴がある。すなわち隠語使用者は、非認知的な影響力を行使しながら、事後的にその行使意図を否認できる余地を持つ。いわば隠語使用は、語に対する訂正過程や語に含まれる脱権力化プロセスそのものが脱白させられるリスクを生じさせるのであり、告発や抵抗が否認されうる危険をより多く含む。逆から言えば、告発された害を発話者自ら「隔離 quarantine」しうるので

許さない) 議論継続を承認し合うための中間領域こそ、多様性の観点から見た個人・社会の発展という目標の基層にあったミルの考えであった。

この結果として、言論を発し、受ける地位が毀損された場合(あるいは毀損するように事実上の権力を行使する場合)に、その格差を是正・回復する促進措置は法がなすべき任務に属する一方で、言論を発し、受ける地位のあらゆる差異の埋め合わせを禁止や規制を通じて実現しようとするのは、法がなすべき任務には属さないこととなる。なぜなら、発信を促進する是正措置の巧拙を競い合わせることは(アーキテクチャによる支援においても)言論を発し、受ける地位を対称にすることを可能にする一方で、禁止や規制を受けるべき属性という主題自体を主たる争点にすることは、言論を発し、受ける地位をも非対称なものに固定するためである。

法による介入が、その内在的制約からごく例外にとどまるのは、この故である。受忍と寛容を涵養する社会と、社会内における共同体間における移動の自由の実質的な行使可能性なくしては、一見したところの消極的自由の保護は名目に終わり、容易に国家による積極的介入へと転化してしまう⁹⁶。ウォルドロンの言葉を借りれば、「政治の状況」を直視しない「アイデンティティの政治の無責任さ」⁹⁷と集合的決定の挫折を、法は任務として加速させる役割を持たないのである。

(2) 訂正可能な空間と人格性 私的自治と民主主義との間

このように、「うち」なる脆弱な個としてのアイデンティティをそのまま「政治の状況」に交渉不能な危害として持ち込むことは、主権者の恣意とともに、「アイデンティティ政治の無責任さ」と集合的決定の挫折に行きつかざるをえない。たとえ運動への動員手法として

ある。

翻って、オンライン上のヘイトスピーチもまた蔑称・隠語と同様に、その文言上の意味内容として、社会におけるマイノリティ成員に対して、彼らの社会的地位を間接的に「普通」以下に引き下ろすことを示唆/顕示する側面に近い言明ではある(沈黙化や抑圧と同様)。オンラインの特性を踏まえれば個別的・集団的に抵抗が可能であるとしても、繰り返される当該言明と社会における「寛容」が、具体的な当該成員の社会における地位を現に固定化する「おそれ」には、社会科学的な蓄積を踏まえれば一定の根拠が認められるかもしれない。その上で、隠語特有の性質をも引き継ぎ、オンライン上のヘイトスピーチは上記の抵抗の際に、弁明としての「そういうつもりはなかった」と否認される危険を有するともいえるだろう。

しかし、こうした権力と抵抗、そして否認の連鎖の中では、特定の発話効果を明示化する作業は、有益ではあれ、規制の許否判定の観点からはほとんど役に立たない。例えば沈黙化の可能性が認められるとしても、沈黙化への抵抗と否認の可能性も、そこには認められるだろうことが示唆されるためである。問題は沈黙化、抵抗、否認が、どれほど肯定的な支持やアーキテクチャ的支援を受け、各々の発話が容易となっているかにある。この支持・支援の社会的・歴史的文脈の特定なしには、沈黙化や抑圧についての哲学的思弁は、ヘイトスピーチの現象的側面の解明という課題を除き、規制の文脈ではほぼ機能を失ってしまう。第3節で見た通り、ある害を除去する規制にはリスクトレードオフが存在し、規制がなす「悪さ」は、その正当性のみならず、その正統性にも及んでいるためである。

⁹⁶ ヘイトスピーチをも取り込む社会の寛容の促進は、異なる「包括的教説」相互間の協同のための基礎構造を、国内におけるコンフリクトとして示すものであると呼べようか。2015年にこの箇所についての報告コメントをした際には明示できなかったが、本稿のベースは同稿にその端緒がある。

⁹⁷ ウォルドロン『ヘイトスピーチという危害』159-162頁にあるが、「ある争点が私のアイデンティティにとって決定的に重要だということによって、私はその争点についての私の見解を政治的に交渉不可能なものとして提示する」ことによる集合的決定の挫折のことを意味する。

有効な瞬間を持つとも、ハードな効果を伴う⁹⁸キャンセルカルチャー群が、特定の属性のみを切り出してそれを包含する同様の動きを攻撃とみなすに至るような場合⁹⁹には、社会における議論継続のための中間領域、いわば個人の取り決めに最も重きを置く私的自治と、集団の取り決めに最も重きを置く民主主義との間に横たわる、取り決め困難な領域を破壊するに至る¹⁰⁰。端的にそこに至った過程に対する責任の所在も、あるいはそこから将来に向けた公正な回復・和解の途も明らかにすることがないためである。

ここから本稿の見立てによれば、抵抗としての表現をエンパワーするとともに、議論継続

⁹⁸ 志田陽子 (2022) 「「表現の自由」をめぐるワインディングロード」現代思想 50 巻 3 号は、まさにこの狭義の意味で「キャンセル」を捉え、契約当事者外からの当事者の解職要求や社会的信用を低下させるような働きかけのことを「キャンセル」と呼ぶ。法的に問題となるキャンセルとしては、本稿もこれに即し、社会的・経済的な地位剥奪としての解雇や契約関係解消、オンライン空間からの締め出しという意味ではフォーラムからのアカウント削除や特定属性への封殺・沈黙化等のハードな形でなされる「キャンセル」という意味で、この語を理解する。なお、こうした立場に対しては、「キャンセル」の意味を広義に取ること許されるキャンセルもそうでないキャンセルもある、と反論されるかもしれない。しかし、こうした概念の混同は（キャンセル現象の記述や例示としては妥当であるとしても）規制をめぐる議論の焦点を逸してしまうだろう。特定の発言や不品位行動に対する非難・批判・説得や、発言者である者との間の友人関係・契約関係からの個別的な離脱といったソフトな「キャンセル」と、そうした行動を他者に要求して集団的に行動することや法を用いた禁止・不可視化措置といったハードな「キャンセル」との間の境界が曖昧化されることは、全くもって当然の帰結ではない。

⁹⁹ 前掲注)41 大屋雄裕 (2020) 「赦しと忘却」アステイオン 93 号では、まさしく黒人差別を訴える”BlackLivesMatter”に対して、“AllLivesMatter”やその他訴えが、BLM に対して敵対的で人種差別的であるとされ、キャンセルされる事態が紹介されている。「すでに二〇一六年の段階から、「すべての命は大切だ」(AllLivesMatter) という表現は BLM に敵対的であり人種差別にあたりと批判されるようになっており、民主党支持者として知られる女優ジュニア・ロベス（彼女自身がプエルトリコ出身の両親のあいだに生まれたラティーノである）でさえ、この表現をハッシュタグとして付したツイートの削除と撤回に追い込まれている。二〇二〇年七月には、「すべての命は大切だ」だけでなく警察官（しばしば青い制服を着ている）の生命保護を訴える「青い命も大切だ」(BlueLivesMatter) や、やはりマイノリティグループと目されている「アイルランド系の命は大切だ」(IrishLivesMatter) といったスローガンを付した商品の販売を取り止めることを、アメリカ最大の（そして世界最大の）スーパーマーケット・チェーンであるウォルマートが表明するに至っている」。

¹⁰⁰ 前掲注 65 大屋雄裕 (2022) 「キャンセル・カルチャーが孕む二つの文脈」『Voice』537 号を参照せよ。もちろん、私的自治領域における一対一の契約内容上の拘束や契約解消が、その効果として何らかの地位からの離脱や利益への期待の剥奪を伴うものであったとしても、違法な拘束・解消あるいはそもそも優越的地位の濫用等による例外を除き、個別的な関係からの離脱や期待剥奪がここでいう「キャンセル」の作用を果たすものとは見做されないだろう。しかしそこから更に踏み込み、当事者の契約には書かれていないある観点からの「不快」あるいは不品位性等の道徳性を理由として、当事者外の者が（批判を超えて）外部から圧力をかけることは、上記の意味での通常の法的な意味での私的自治の範疇の限界点に位置していることもまた明らかである。なぜなら、契約関係であればそれによる履行請求や解消の妥当性、あるいはその地位上の格差（例えば優越的地位の濫用法理の適用可否）が、法的な意味で終局的には対等な道具立ての下での解決を志向する司法的解決に付されることが期待されるのに対し、外部からの匿名的圧力の集積は、自らの利害に直接には関わらない場面において（例えば当該人物と雇用先との契約主体を介して）「キャンセル」の結果を間接的に引き起こしうるにもかかわらず、その（比喩的には法の下に出頭すべき）責任主体性を曖昧化したまま、当該結果に対する責任を取ることを避け、あるいは結果をめぐる回復の営みに参与せず、事後的に自らの行為と結果との間の連関を機会主義的に「否認」しうるためである。

比較対象として、例えば消費者集団が自らもまたその商品の利益を享受する主体であることを理由とし、その自律的意思を集積することを呼びかける不買運動や、社会的な誉れを総合的に受けることが期待される伝統的な賞の付与や銅像建造において特定の不品位を排除する営為を横に置いて検討してみよう。これらと上記の「キャンセル」とが異なるのは、正しくこの責任主体性に比例的であることがわかるだろう。不買運動においては「買う」主体は自らが対価を支払ってきた、あるいは支払うだろう契約関係にあり、法的な意味で把握される私的自治の範疇に属する。翻って伝統的な賞からの不品位な人格・表現の排除や銅像となったかつての偉人の評価替えに由来する引き倒しは、まさにその「伝統」の価値（あるいは価値変動）が共同体に属する一人一人に共有された価値（あるいは価値変動）として潜在的に含有されていることに求められよう。

を可能とする促進的環境形成のために、立法・行政・司法作用にどのように権限配分するか、（プロバイダ事業者やプラットフォーム等）私人を通じた法執行として許否するかこそが、法を用いてヘイトスピーチへの対応をなす際の実質的な中心課題となる。切り分けと「キャンセル」は、事実上、発言と発言者を結びつけるとともにその差異の発露を不可視化する¹⁰¹ことで、発言者が社会の中で自らの誤りの「おそれ」に開かれ、自らの訂正を施しながら進む可能性を失わせる。単に欲望を地下化させ、増長させるのみならず、地上において自他の差異のただ中で、自らを訂正することが可能な市民の尊厳ある一人として遇する努力を無化してしまうだろう。

もちろん、こうした訂正を待つことへの懸念にも、全く理由がないわけではない。害が顕在化しているケースはもとよりとして、時に一般的な意味での表現の自由の「武器化」が論じられることがあるのもその一例に挙げられよう¹⁰²。文化闘争場面において保守的とされる側が、表現の自由の主張を盾に自らの政治主張を行い、裁判所が受け入れることが、この「武器化」の達成であるとされる。確かに、歴史問題の司法化¹⁰³の功罪も盛んに論じられる現在、「正しさ」を決する場が裁判であるという臆見（やそうしたアピール）が通用する場面において、一見したところの表現の自由の「正当」な行使に、様々な暴力性への加担を見てとることは（それら議論が法学上妥当か否かは措いておくが）不可能ではない。

しかし、対等な道具立てで解釈を争ったにもかかわらず、結果として特定の立場への「不都合」な帰結をもって望ましからざる法的概念の「武器化」と評する限りにおいては、少なくとも法的な議論においてはこの武器化概念がもとより転倒していると評価せざるをえない。先に挙げた歴史問題の司法化批判が司法における「武器」となることへの懸念やその除去要求ではなく、そもそも歴史形成の場の中心は司法ではなく、歴史形成に不断に関与するのは社会の役割であると強調していたことが、まさにここで想起しうる。裁判もまた法制度の中で正統性を分有する一つの機関に過ぎず、社会における「正しさ」を決する場ではない。新たな裁判にも、また裁判外にも議論のフォーラムは開かれている。したがって、この「武器化」を嘆く言説こそ、抵抗をも可能とする表現の自由を支えるエンパワメントと促進的環境形成に向けられた議論継続とを信頼することなく、「表現の自由」を我々の社会の中に位置づけることを「おそれ」ているとさえいえよう。

ウォルドロンの見解の検討から把握できるように、表現の自由が十全に保障された空間を求めるとは、自らの表現が聞き取られる価値があるものとして受け止められる空間を求

¹⁰¹ 成原慧（2023）「キャンセルカルチャーと法」ではこのキャンセルの機能を、キャンセル要求が事実上果たすとしても基本的には表現の自由の枠内に属し、保護に値するとしていた。本稿の観点からは、仮にこれが認められる場合に、個人に充てられた侮蔑・名誉毀損等に該当しない、直接的でないプロパガンダ的ヘイトスピーチが「機能的」にキャンセル作用を果たす場合にも同様の帰結に導かれるものと思われるが、成原がそのような帰結をよしとするかは定かではないように思われる。備忘的に疑問としてここに記す。

¹⁰² 森口千弘（2022）「社会の分断がもたらす人権の「武器化」——マイノリティの権利の観点から」『〈分断〉と憲法・政治・社会から考える』

¹⁰³ 武井彩佳（2021）『歴史修正主義-ヒトラー賛美、ホロコースト否定論から法規制まで』および同氏（2022）「歴史否定論と陰謀論」世界 961号 130-140頁を参照せよ。

めることでも、他者の感情が毀損されうる「おそれ」を欠いた安全な空間を求めることではない。表現の対称性に鑑みればむしろ反対に、自らを傷つけるかもしれない相手の表現に向き合う空間を求めることであり、あるいは向き合うことが困難な事情があるならば、自らの代わりに自らを支持・支援する他者の「より多くの」表現が響き渡る空間を求めることである。誰もが「おそれ」を抱きつつも、「おそれ」を排除することなく、「正しくおそれる」ためには、他者への批判が妥当し、あるいはその可能性が残存する限り、自らを訂正する可能性にも開かれる必要がある。

(3) 「おそれ」を正確に恐れうる環境

時に J. S. ミルは言論の自由の核心を真理の発見に見たと要約されることがある。しかしより正確には、その時々において通用する一般的見解とは異なる意見が一般的見解へと挑戦する議論継続の有益性の内に見た、という方が正確である。これは二つの解釈を持つ。第一には文字通りに、その時に正しいものとして通用している意見が誤っている可能性に開かれ（真理発見）、意見の正しさを反対論に対抗することで明瞭に理解する（真理確証）側面のみならず、異なる意見がそれぞれに正しさを分有している場合にその両者を保存する（真理保存）側面に、その核心たる所以を持つというものである¹⁰⁴。しかしミルが同じくらいに重きを置いていたのは、そうした真理をめぐる発見・確証・保存の営為が特定の科学者や思想家の占有物ではなく、「普通の間を彼らの到達できる限りの精神的高度に到達することを得せしめる」¹⁰⁵ために必須の要件であり、それこそが言論の自由の核心たる所以を持つというものであった。

この内、前者は社会功利に通じる真理の側面として主に『自由論』第二章で論じられる「真理に基づく正当化」に、後者の自己開発的側面は主に『自由論』第四章で論じられる「人格に基づく正当化」に区分されることがある¹⁰⁶が、上記引用箇所近接性が示すように、議論を引き受け、立てることは、この両方を横断している。自身のイニシアティブで意見が提起しえ、議論が起こらうという「信頼」は、挑戦された一般的見解を擁護する者に合理的根拠とともに理解する契機を与えるとともに、強いられた教説としてではなく市民各人の人格の発露を通じた確信を成長させる助けとなる。したがって、ミルによれば、社会に対する真実を巡る挑戦はまずもって（挑戦を受けた者が矢面に立つか否かは別として）議論に参加することで受け止め、結果としてその度ごとに真実は訂正と確証に開かれねばならないことに帰結しよう。

社会における平等な地位へと疑問符が付され、あるターゲット集団に対して自尊心を蝕む「おそれ」が蔓延したときに、被ターゲット集団及び非ターゲット集団の間で義務付けられることとは何か。本稿の検討によれば、社会における防御として、ある種の表現をスケー

¹⁰⁴ 前掲注)82 ミル (1971) 『自由論』 93-94 頁

¹⁰⁵ 前掲注)82 ミル (1971) 『自由論』 71 頁

¹⁰⁶ マーサ・ヌスバウム (2010) 『感情と法』(慶應義塾大学出版界会、河野哲也 監訳) 第7章 第二節 409 頁

プゴートに捧げ、刈り取る営為ではなく、むしろ「おそれ」のある表現の受け手をエンパワメントし、「安心」を再構築する努力が、市民としての上記集団には義務付けられるはずである。発話による害の「おそれ」が、単なる不快を超えて実際の「害」となるか否かは、まさしくそうした「おそれ」を取り巻く市民 citizen 相互でその「おそれ」を受け止めうるかに依存したことを思い起こすべきであろう。

ウォルドロンは、著名な『ヘイトスピーチという危害』を著した翌年の2013年、別の論文において、人間の尊厳とシティズンシップ citizenship の尊厳を並行的に論じていた¹⁰⁷。シティズンシップは、それぞれの政治体における、特別であり、かつ、平等で広範な社会的地位という、一見すると相反する要素を含む。シティズンシップは歴史を通じて変化し、場所によっても異なるが、いかなるシティズンシップもひとかたまりの権利・権力・義務・責任から成り立ち、その内容や適用は法的問題として定められる点では共通する。シティズンシップの尊厳は、特権階級の特別な権利ではなく万人への普及が目指されるが、それは同時に、自らの利益を超え、全員の利益に公正に応える政策・法律・枠組みに注意を向ける義務を引き受ける人に向けた尊敬を含む。すなわち、シティズンシップが尊厳たるのは、他者の尊厳を尊重する義務、法に服従し政治的敗北を受け入れる義務、市民権を能動的に行使する責任に結びついているためである¹⁰⁸。

責任と結びつくそのエンパワメントは、単にマイノリティ成員を脆弱的存在として遇することによって社会から危険をあらかじめ除去し、あるいは成員を危険を拭えないという理由で社会的営為から「隔離」し、保護されるべき存在に留めおくこととは異なる。たとえある意味での保護に向けられた「隔離」であっても、家族集団や宗教的集団内部への近時の法的介入（DV等）の進展を背景に、その部分社会性の限界を示すものであることは認識されているだろうし、それゆえに共同体間における移動の自由の実質的な行使可能性が鋭く問われていることはその証左である。逆もまた然りであり、「おそれ」のある表現の表現者側でさえも、刑罰を用いて物理的に（あるいはコンテンツモデレーションを介して空間的に）「隔離」することは、むしろ社会復帰を期し、社会内処遇の必要性との緊張関係に立っている¹⁰⁹。ある属性に基づくマイノリティ成員は（他者からのスティグマの帰結であれ、自己スティグマ¹¹⁰の帰結であれ）一方的に保護されるべき「脆弱」な主体ではなく、むしろそ

¹⁰⁷ ジェレミー・ウォルドロン（2017）「シティズンシップと尊厳」思想 1114 巻。なおウォルドロンによる 2008-2009 タナー講義“Dignity, Rank, and Rights”も同論点を含む。前掲注）

¹⁰⁸ 同上 125-127 頁を参照せよ。この二重性を、憲法学者の蟻川恒生は、たとえ文化的少数派に属する場合であろうとも、「自らの文化的アイデンティティを公共空間の中で主張し貫徹する」ことを自ら抑制することができる個人でありつつ、他者に対しても「抑制するよう要求することができる個人」であるとする。蟻川恒生「憲法学に「個人」像は必要か」『尊厳と身分、憲法的思惟と「日本」という問題（2016年）248頁を参照せよ。

¹⁰⁹ OECD のメンタルヘルス関連資料における長期入院者数に見られるように、本文に記した社会内処遇の課題は刑事的な課題に限られず、「脱施設化」や共生社会一般の問題として理解される。本邦における精神疾患の長期入院者数に見られる精神疾患を抱えた者の社会からの「隔離」実践が、社会の無理解や拒絶心によって正当化されないように、社会における「おそれ」はその発話者の「隔離」を当然には正当化しないのである。OECD 資料として、下記を参照せよ。 <https://www.oecd.org/els/health-systems/mental-health.htm>

¹¹⁰ マーサ・ヌスバウム「老いとスティグマと嫌悪感」（田中あや訳）16頁

れを「脆弱」化させてきた正常性・完全性に基づく社会秩序¹¹¹を変更することによって、相互協力し合う社会のメンバーとなることが期待されるだろう。

(4) リスクのキャンセルから、リスク含みの選択のエンパワメントへ

市民を、等しい市民として尊重するためには、多様な市民の存在と各々が信じる（人生についての理に適った）包括的教説を多元的に保障する環境が必要である。マーサ・ヌスバウムは、その著書『感情と法』において「社会を人間の尊厳を守る場所にする役割を果たし、市民が恥辱とスティグマにさらされない生活を送れる」¹¹²環境を、促進的環境と呼んでいたことがここで想起される。貧困、差別・ヘイトクライム、性的マイノリティ、障害（ディサビリティ）の諸例に挙げつつ、ヌスバウムは「正常」という規格によって社会をデザインしたことで市民の一部を「脆弱」な者とのスティグマへと押し込める社会秩序は、秩序だった decent ものではなく、その是正のコストがリベラルな社会には要求されると主張していた。必要なのは（社会的専制と同様に、完全性の専制¹¹³を排し）「さまざまな障害を負った人が「巧妙な相互作用」の人生を生きることができるよう「行動を容易にする環境」を創造する」¹¹⁴ことである。

このように、「隔離」から自律的な離脱/再参入へとシフトしつつある法の力点移動を踏まえるならば、「おそれ」に満ちた中でも新たなコミュニケーションに参入するための「信頼」を構築することに、法の営為もまた向けられるはずである。逆説的なことに、尊厳に満ちた社会を動揺させる「おそれ」をもつヘイトスピーチに刑罰を付与し、法で規制しなければならない場面とは、シティズンシップの尊厳を普及することを止めた、飛び地としての「二級市民」を許容する社会においてであることに帰結することになる。

確かに被害は、社会に埋め込まれた特定のアイデンティティ構成的属性への毀損を介し、局所化されて生じうる。まさにその場面では、声を上げる人を具体的に一時的には保護し、承認を回復することが求められるかもしれない。しかし、法が提供するコントロールには、保護・承認に向けて論争を回避することを含む「キャンセル」作用を果たすことは、例外的にしか含まれない。市民一人ひとりの脆弱さと相互依存の上でなお、潜在的には常に生じうる論争的な表現とともに立ち向かう、連帯とエンパワメントによって、シティズンシップの尊厳を各アクターが発揮する環境形成に、法のコントロールは向けられている。少なくとも現実の属性を直接に反映することがないオンライン上の表現空間においては、アイデンティティ構成的属性へと向けられたヘイトスピーチとその害を抽象化したまま規制の議論を

¹¹¹ 前掲注)103 マーサ・ヌスバウム『感情と法』424頁を参照せよ。ミルは社会的専制を、生活の細部に入り込むことで、逃れることを不可能にする魂の奴隷化として描き出し、考え方や慣習についての行動ルールの実事上の強制及び個性の発揮の除去の双方へと通じるものとしていた。これに対しヌスバウムは完全性の専制を、「主流を成す社会規範が、その規範と合致しない生き方に対し、法を通じて与える壊滅的な影響」として描いている。

¹¹² 同上 355 頁

¹¹³ 同上 386 頁

¹¹⁴ 同上 399 頁

行うことなく、複数のアイデンティティを構築しつつある個別の人格を、「尊厳」と「自尊心」に満ちた、複数的な存在へと変じるプロセスへと開くことが求められるであろう。我々は、完全ではない世界において、互いを損なうかもしれないリスクとともに、しかしそのリスクに晒された者を助け、癒し、またリスク含みかもしれない他者を許す中で、被害者と加害者の区別を乗り越えてともに生きていくしかない存在でしかないし、しかしだからこそ、そういう存在でありうるという「信頼」を構築する努力が課されているはずである。

おわりにかえて 論考[展開]に向けて

縷縷述べてきたように、ヘイトスピーチをめぐる状況は、理論的にも社会的・歴史的文脈に規定されざるを得ない側面があるとともに、実践的にも引きずられる側面が多くある。ヘイトスピーチをめぐる状況には、エビデンスの過剰利用の危険が常につきまとうのである。それが感情的なバックラッシュの一因となることも想定しやすい。こうしたオンライン空間における尊厳促進的環境としての法と、規制による「感情」のわだかまりを解きほぐすのは困難ではある。その中で、本稿ではヘイトスピーチの事例からオンライン上の（隠喩的）ゾーニング／キャンセル文化を考えてきた。そして、アイデンティティ構成的属性について、単一的帰属と複層的構成とを区別することで、後者に対するオンラインヘイトスピーチの害の適応段階における限定性を見てきた。

しかし、論考[展開]でも見ていくように「シティズンシップの尊厳」を潜在的にもつものとは、武装解除を求めるのではなく、武器にも耐えられる「タフ」な空間を求め、自他の訂正可能性に開かれつつそれを絶えず確証し続ける存在である、というのが本稿から見てとれる展望である。しかし、果たしてこれは将来における断罪の日が来ることを受け入れつつも、将来における赦しを期待する、寛容に満ちた社会の像だろうか¹¹⁵。本稿で記述してきた各種のリスクトレードオフのただ中で「法を用いてものごとをなす」法（哲）学から見たヘイトスピーチ規制の第一の指導理念は、安心と自尊心、そしてシティズンシップの尊厳を育むだろう、ロバート・ブランダム曰くの「信頼の精神」¹¹⁶であり続けられるだろうか。

以上、紙幅の都合上、（法）規制を問う前提を整理することまでが本稿の対象となったものの、本稿自体への訂正を期しつつ、別稿へと繋げて一旦筆をおく。

以上

¹¹⁵ Joseph Raz (1986), *Morality of Freedom*, Oxford pp396-398 を参照せよ。

¹¹⁶ Robert Brandom(2019), *A Spirit of Trust: A Reading of Hegel's Phenomenology* 及び Gilles Bouché (ed.)(2020), *Reading Brandom On A Spirit of Trust*