

Nordic Lectures in Pragmatism 3

**Robert Brandom**  
Analytic Pragmatism, Expressivism  
and Modality

**PNordic  
Pragmatism  
Network**

Helsinki, Finland  
30 September –  
2 October 2014

[www.nordprag.org](http://www.nordprag.org)

Photo: © 2014  
"Nordic Pragmatism" by Paul K.

# Analytic Pragmatism, Expressivism, and Modality

## **The 2014 Nordic Pragmatism Lectures**

Robert B. Brandom

Robert Brandom's recent work aims to reconceive and renew pragmatist thought by on the one hand reconnecting it with its roots in the German Idealist tradition, and on the other hand synthesizing it with analytic philosophy of language, epistemology, and metaphysics. These lectures provide an overview of how he proposes to do that. A central figure in this reconceptualization of pragmatism and its history is the brilliant and notoriously difficult mid-century American philosopher Wilfrid Sellars. Brandom explains how Sellars's ideas can form the basis for contemporary analytic pragmatist reworkings of the Kantian idea of categories, the significance of modal vocabulary, and the metaphysics of universals. The lectures are drawn from Brandom's most recent books *Perspectives on Pragmatism: Classical, Recent, and Contemporary* [Harvard University Press, 2011] and *From Empiricism to Expressivism: Brandom Reads Sellars* [Harvard University Press, forthcoming in 2014].

Robert Brandom is Distinguished Professor of Philosophy at the University of Pittsburgh. He is a fellow of the American Academy of Arts and Sciences, and the recipient of an A.W. Mellon Foundation Distinguished Achievement in the Humanities Award. He has delivered the John Locke lectures at Oxford (published as *Between Saying and Doing: Towards an Analytic Pragmatism* [Oxford University Press, 2008]), the Hempel lectures at Princeton, the Townsend lectures at Berkeley, and the Woodbridge lectures at Columbia. He has held visiting fellowships at All Souls College Oxford and the Center for Advanced Study in the Behavioral Sciences at Stanford, and was Leibniz Professor at the University of Leipzig. He is the author of twelve books, including *Making It Explicit* [Harvard University Press, 1994], which have been translated into many languages.

## **Analytic Pragmatism, Expressivism, and Modality**

Lecture One: From German Idealism to American Pragmatism, and Back

Lecture Two: On the Way to a Pragmatist Theory of the Categories

Lecture Three: Towards an Analytic Pragmatism: Meaning-Use Analysis

Lecture Four: Pragmatism, Inferentialism, and Modality  
in Sellars's Arguments Against Empiricism

Lecture Five: Modal Expressivism and Modal Realism: Together Again

Lecture Six: Sellars's Pragmatic Metalinguistic Expressive Nominalism

## 2014 Nordic Pragmatism Lectures:

### Analytic Pragmatism, Expressivism, and Modality

#### Lecture 1

#### From German Idealism to American Pragmatism—and Back

##### I. Kant and Hegel

Developments over the past four decades have secured Immanuel Kant's status as being for contemporary philosophers what the sea was for Swinburne: the great, gray mother of us all. And Kant mattered as much for the classical American pragmatists as he does for us today. But we look back at that sepia-toned age across an extended period during which Anglophone philosophy largely wrote Kant out of its canon. The founding ideology of Bertrand Russell and G.E. Moore, articulating the rationale and fighting faith for the rising tide of analytic philosophy, was forged in a recoil from the perceived defects of a British idealism inspired by Hegel. Mindful of the massive debt evidently and self-avowedly owed by Hegel to Kant, and putting aside neo-Kantian readings of Kant as an empiricist philosopher of science that cast him in a light they would have found more favorable, Russell and Moore diagnosed the idealist rot as having set in already with Kant. For them, and for many of their followers down through the years, the progressive current in philosophy should be seen to have run directly from Locke, Leibniz, and Hume, to Mill and Frege, without any dangerous diversion into the oxbow of German idealism.

What did the pragmatists learn from Kant? I want to focus on two of Kant's master ideas: what I'll call his *normative turn*, and what I'll call (tendentiously but only proleptically) his *pragmatist methodology*. I think that we should still care today about these ideas—ideas which were for complicated reasons largely invisible to classical analytic philosophy. As I

understand his work, Kant's most basic idea, the axis around which all his thought turns, is that what distinguishes exercises of judgment and intentional agency from the performances of merely natural creatures is that judgments and actions are subject to distinctive kinds of *normative* assessment. Judgments and actions are things we are in a distinctive sense *responsible* for. They are a kind of *commitment* we undertake. Kant understands judging and acting as applying *rules*, concepts, that determine *what* the subject becomes committed to and responsible for by applying them. Applying concepts theoretically in judgment and practically in action binds the concept user, commits her, makes her responsible, by opening her up to normative assessment according to the rules she has made herself subject to.

The responsibility one undertakes by applying a concept is a task responsibility: a commitment to *do* something. On the theoretical side, what one is committed to doing, what one becomes liable to assessment as to one's success at doing, is integrating one's judgments into a whole that exhibits a distinctive kind of unity: the synthetic unity of apperception. It is a systematic, rational unity, dynamically created and sustained by drawing inferential consequences from and finding reasons for one's judgments, and rejecting commitments incompatible with those one has undertaken. Apperceiving, the characteristically sapient sort of awareness, is discursive (that is, conceptual) awareness. For it consists in integrating judgments into a unity structured by relations of what judgments provide reasons for and against what others. And those rational relations among judgments are determined by the rules, that is the concepts, one binds oneself by in making the judgments. Each new episode of experience, paradigmatically the making of a perceptual judgment, requires integration into, and hence transformation of the antecedent constellation of commitments. New incompatibilities can arise, which must be dealt with critically by rejecting or modifying prior commitments. New joint consequences can ensue, which must be acknowledged or rejected. The process by which the whole evolves and develops systematically is a paradigmatically rational one, structured by the rhythm of inhalation or amplification by acknowledging new commitments and extracting new consequences, and exhalation or criticism by rejecting or adjusting old commitments in the light of their rational relations to the new ones.

Kant's new normative conception of what the activity of judging consists in, of what one must be *doing* in order to be judging (a corresponding story applies to acting), puts important

structural constraints on how he understands the judgeable *contents* for which one is taking responsibility in judgment. The dominant order of logical and semantic explanation of the tradition Kant inherited began with a doctrine of terms or concepts. On that base, a doctrine of judgments was erected, and then finally a doctrine of consequences or syllogisms. But the minimal unit of *responsibility* is the judgment. It is judgments, not concepts, that one can invest one's authority in, commit oneself to, by integrating them into an evolving constellation that exhibits the rational synthetic unity of apperception. Accordingly, in a radical break with his predecessors, Kant takes judgments to be the minimal units of awareness and experience. Concepts are to be understood analytically, as functions of judgment—that is, in terms of the contribution they make to judgeable contents. To be candidates for synthesis into a system exhibiting the rational unity characteristic of apperception, judgments must stand to one another in relations of material consequence and incompatibility. So if one is to understand judging also as the application of concepts, the first question one must ask about the contents of those concepts how the use of one or another concept affects those rational relations among the judgeable contents that result. This methodological inversion is Kant's commitment to the *explanatory primacy of the propositional*. It is a methodological commitment that will be seconded by Frege, whose *Begriffsschrift* is structured by the observation that it is only judgeable contents to which pragmatic force can attach, and by Wittgenstein, who in the *Investigations* gives pride of place to sentences as the only kind of linguistic expression that can be used to make a move in a language game.

Kant's thought here, I think, is that alongside the *local* order of explanation, which looks to the contents of the particular concepts applied in judging to explain the specific possibilities of rational integration of judgeable contents containing them (their inferential grounds, consequences, and incompatibilities), there is a *global* order of explanation according to which one must understand what conceptual content *is* in terms of what judgeable contents are, and must understand that in terms of what one is *doing* in judging, in making oneself responsible for such contents. The *functionalism* about conceptual contents that consists in understanding them as functions of judgment, which is the practical expression of methodological commitment to the explanatory primacy of the propositional, is motivated by an overarching methodological



*pragmatism* according to which semantics must answer to pragmatics (in a broad sense).<sup>1</sup> It is the strategy of understanding discursive *content* in terms of what one is *doing* in endorsing or applying it, of approaching the notions of judgeable, and therefore conceptual content generally, in terms of the constraints put on it by requirement derived from the account of the *activity* of judging.

Though I have for expository reasons focused my sketch on the cognitive, theoretical side of Kant's thought, it is important to be clear that pragmatism in the sense I am attributing to Kant is *not* a matter of giving explanatory priority to the practical over the theoretical, to exercises of agency over exercises of cognition. Rather, within both the practical and the theoretical spheres, it is understanding *content* in terms of *force* (in Frege's sense): what is judged, believed, or done in terms of one must *do*, what activity one must engage in, to be judging, believing, or doing it. Kant, I am claiming, should be thought of as a pragmatist *avant la lettre* because of the way his normative theory of conceptual *activity* (theoretical and practical) shapes his account of conceptual *content* (both theoretical and practical).

I read Hegel as taking over from Kant commitment both to a normative account of conceptual doings, and to a broadly pragmatist approach to understanding the contents of our cognitive and practical commitments in terms of what we are doing in undertaking those commitments. I see him as taking an important step toward *naturalizing* the picture of conceptual norms by taking those norms to be instituted by public social recognitive practices. Further, Hegel tells a story about how the very same practice of rational integration of commitments undertaken by applying concepts that is the synthesis at once of recognized and recognizing individual subjects and of their recognitive communities is at the same time the historical process by which the norms that articulate the contents of the concepts applied are instituted, determined, and developed. He calls that on-going social, historical process "experience" (Erfahrung), and no longer sees it as taking place principally between the ears of an individual. Closer to our own time, we have seen a version of this development repeated as Quine, in a pragmatist spirit, rejects Carnap's two-phase picture, according to which first one institutes meanings (a language) and then applies them (adopts a theory), in favor of a unitary

---

<sup>1</sup> Later on (in Section V) I will suggest a somewhat narrower use of the term "methodological pragmatism".

process and practice of using expressions that must be intelligible at every stage as settling (insofar as it is settled) both what one means and how one takes things to be.

## II. Classical American Pragmatism

In the broadest terms, the classical American pragmatists, Peirce, James, and Dewey, developed this German idealist tradition by completing the process of *naturalizing* it, which had begun already with Hegel. In their hands, it was to take on the shape of an empirical scientific account of us and our transactions with our environment. The sort of understanding they sought was decisively shaped by two new models of scientific explanation, codifying new forms of intelligibility characteristic of late nineteenth century science. Principal among these, of course, was Darwinian evolutionary explanations. The other form of explanation that was coming to maturity in the science of the day was statistical explanation. Pragmatism begins with a philosophy of science, pioneered by Peirce, that saw these two explanatory innovations as aspects of one conceptual revolution in science.

One dimension along which evolutionary and statistical explanations differ from those of the older mathematical physics concerns the dominant *modality* in which they are expressed. The modality of Newtonian laws is *necessity*. One explains something by showing that it is necessitated by eternal, exceptionless, universal laws. Evolutionary and statistical explanations explain *contingent* happenings, by displaying conditions under which they can be seen to have been *probable*. Both are ways of making intelligible the contingent emergence of collective order from individual randomness.

The original subject-matter of evolutionary explanations was, of course, the process by which biological species arise and diversify. Taking his cue from the way in which statistical explanation had been generalized from its original applications in social science to provide the basis for the triumph of thermodynamics in physics, Peirce substantially generalized evolutionary-statistical forms of intelligibility in two different directions. Most important was an

idea that was picked up and developed by James and above all by Dewey: the recognition that *evolution*, at the level of species, and *learning*, at the level of individuals, share a common *selectional* structure. Both can be understood as processes of *adaptation*, in which interaction with the environment preserves and reproduces (selects) some elements, while eliminating others. This insight is encapsulated in the concept of habit, and the picture of individual learning as the evolution-by-selection of a population of habits. This master idea made possible the naturalistic construal of a cognitive continuum that runs from the skillful coping of the competent predator, through the practical intelligence of primitive hominids, down to the traditional practices and common sense of civilized humans, all the way to the most sophisticated theorizing of contemporary scientists. All are seen as of a piece with, intelligible in the same general terms as, biological evolution.

The other direction in which Peirce generalized the evolutionary statistical selectional model of explanation was to *inorganic* nature. What those older scientific naturalists, for whom the paradigm of scientific understanding was Newtonian physics rather than Darwinian biology, had taken to be eternal, immutable, necessary, universal laws of nature, Peirce now sees as themselves in the largest sense “habits” of the universe—a kind of order that has arisen contingently, but ultimately statistically explicably, by a selectional-adaptational process operating on a population of such regularities, which in turn provides the dynamic habitat to which all must collectively adapt. There is no guarantee that any such accommodation will succeed permanently. As with habits learned by individuals, some of the lawlike regularities may prove more robust and others more fragile. The older picture of laws shows up as at best only approximately true, an idealization extrapolating a situation that actuality approaches at most asymptotically.<sup>2</sup> The naturalism of the classical American pragmatists was shaped by the new sort of nature they had been taught about by the best science of their times—a nature viewed through the lens of the new forms of statistical and selectional explanation.

The pragmatists’ new form of *naturalism* was coupled with a new form of *empiricism*. The experimental scientific method is seen as just the explicit, principled distillation of the selectional learning process that is the practical form common to intelligent creatures at all stages of development. Dewey’s term for that process, in all its varieties, is ‘experience’—the axial

---

<sup>2</sup> James endorses this Peircean idea in Lecture II of *Pragmatism*. [ref.]

concept of such central works as *Experience and Nature* and *Art as Experience*. (So central is the concept to Dewey's thought that sometimes in reading these works it is difficult to overcome the impression that he is, as Rorty once put it, "using the term 'experience' just as an incantatory device to blur every conceivable distinction.") Experience in this sense is not the ignition of some internal Cartesian light—the occurrence of a self-intimating event of pure awareness, transparent and incorrigible to the subject of the experience. Experience is *work*: the application of force through distance. It is something *done* rather than something that merely *happens*—a process, engaging in a practice, the exercise of abilities, rather than an episode. It is experience, not in the sense of *Erlebnis* (or *Empfindung*), but of Hegel's *Erfahrung*. It is the decidedly non-Cartesian sense of 'experience' in which a want-ad can specify "No experience necessary," without intending thereby to invite applications from zombies. Earlier empiricists had thought of experience as the occurrence of conscious episodes that provide the raw materials for learning, via processes such as association, comparison, and abstraction. For the pragmatists, experience is not an input to the learning process. It just *is* learning: the process of perception and performance, followed by perception and assessment of the results of the performance, and then further performance, exhibiting the iterative, adaptive, conditional-branching structure of a Test-Operate-Test-Exit loop. The result of experience is not best thought of as the possession of items of *knowledge*, but as a kind of practical *understanding*, a kind of adaptive attunement to the environment, the development of habits apt for successful coping with contingencies. It is knowing how rather than knowing that.

Ontological naturalism and epistemological empiricism are both encouraged by the idea that the rise of modern science, the most successful social institution of the past three hundred years, can teach philosophers the most important lessons both about how things are and how we can understand them. But from the beginning they have typically stood in significant tension with one another. The furniture of Newton's natural world does not include Locke's mind. And Hume can find nothing in experience by which we could come to know or understand laws such as Newton's as having the necessity that distinguishes laws from mere regularities. Nor is this tension a characteristic only of Enlightenment naturalism and empiricism. It equally afflicts the twentieth-century versions. The two principal wings of the Vienna Circle, which Carnap struggled heroically to keep from flying off in different directions, were distinguished precisely by their answers to the question: when empiricism and naturalism conflict, which should be

relaxed or given up? Schlick urged the preeminence of empiricism, while Neurath was committed to the priority of naturalism. Quine never fully reconciled his (logical) empiricist hostility to modality with his naturalist privileging of the deliverances of science.

The classical pragmatist versions of naturalism and empiricism, though, fit together much better than the versions that preceded and succeeded them. Far from being in tension, they complement and mutually support one another. Both the world and our knowledge of it are construed on a single model: as mutable, contingent products of statistical selectional-adaptational processes that allow order to pop to the surface and float in a sea of random variability. Both nature and experience are to be understood in terms of the processes by which relatively stable constellations of habits arise and sustain themselves through their interactions with an environment that includes a population of competing habits. There is no problem in principle in finding a place for experience construed as learning in nature construed as evolving. Nor is there any analog of the traditional complementary problem of understanding how experience construed as the dynamic evolution of habits can give its subjects access to the modally robust habits of the things those knowers-and-agents interact with, adapt, and adapt to. The pragmatist forms of naturalism and empiricism are two sides of one coin.

The pragmatists' conception of experience is recognizably a naturalized version of the rational process of critically winnowing and actively extrapolating commitments, according to the material incompatibility and consequence relations they stand in to one another, that Kant describes as producing and exhibiting the distinctive synthetic unity of apperception. For that developmental process, too, is selectional (though not statistical). Some commitments (theoretical and practical) thrive and persist, in concert with their fellows, while others are modified or rejected as unable to flourish in that environment. It might be thought fanciful to focus on this common structure in light of the substantial difference between the conceptions: Kant's process is structured by *rational*, conceptual relations of incompatibility and consequence, while the pragmatists' version is structured by *natural*, causal relations of incompatibility and consequence.

But the pragmatists would disagree. For they introduce not only a new conception of experience, but also a new conception of reason. They understand the rationality of the theoretical physicist as continuous with the intelligence of the culturally primitive hunter and the

skill of the non-human predator. The grooming and development of discursive cognitive and practical commitments is a learning process of a piece and sharing a structure with the achievement of practical attunement to an environment and the acquisition of habits successful in that environment that in one form or another is a part of the natural history of all sentient organisms. Reason and intelligence in this sense can be seen (albeit in an inflexible and unlearned form) already in the maintenance of an equilibrium by that emblem of the industrial revolution: the fly-wheel governor. The nature of the pragmatists is through and through a rational nature—not just the part of it that is intelligible as experience.

### III. Fundamental Pragmatism

The more specific strategy by which the classical American pragmatists sought to naturalize the concept of experience—to demystify and domesticate it, to disentangle it from two centuries of Cartesian encumbrances—is what I will call *fundamental* pragmatism. This is the idea that one should understand knowing *that* as a kind of knowing *how* (to put it in Rylean terms). That is, believing *that* things are thus-and-so is to be understood in terms of practical abilities to *do* something. Dewey, in particular, saw the whole philosophical tradition down to his time as permeated by a kind of platonism or intellectualism that saw a rule or principle, something that is or could be made conceptually or propositionally explicit, behind every bit of skillful practice. He contrasted that approach with the contrary pragmatist approach, which emphasizes the implicit context of practices and practical abilities that forms the necessary background against which alone states and performances are intelligible as explicitly contentful believings and judgments. In this reversal of the traditional order of explanation, Dewey is joined by the Heidegger of *Being and Time*, with his project of understanding *Vorhandenheit* as a precipitate of the more ‘primordial’ *Zuhandenheit*, and by the later Wittgenstein. All three thinkers are downstream from Kant’s fundamental insight about the normative character of cognition and agency, and share a commitment to the explanatory priority of norms implicit as proprieties of practice to norms explicit as rules or principles.

I mean the rubric “fundamental pragmatism” to be a relatively loose and elastic description, whose parameters can be adjusted or interpreted so as to fit the methodology of many thinkers, who might differ in many other ways. It is supposed, for instance, to include both the order of explanation that lead Quine to criticize “myth of the museum” in thinking about meaning and that Sellars employs in criticizing the “myth of the given” in thinking about sensory experience. It depends on a contrast, which may be filled-in in different ways, between something on the implicit, know-how, skill, practical ability, practice side and something on the explicit, conceptual, rule, principle, representation side. So we might distinguish between two grades of intentionality: practical and discursive. Practical intentionality is the kind of attunement to their environment that intelligent nonlinguistic animals display—the way they can practically take or treat things as prey or predator, food, sexual partner or rival and cope with them accordingly. Discursive intentionality is using concepts in judgment and intentional action, being able explicitly to take things to be thus-and-so, to entertain and evaluate propositions, formulate rules and principles. The fundamental pragmatist aspiration is to be able to exhibit discursive intentionality as a distinctive kind of practical intentionality. This project can take a strong reductionist form. For instance, what I have elsewhere<sup>3</sup> called the “pragmatist version of artificial intelligence” claims that there is a set of practices or abilities that are non-discursive, in the sense that each of them can be engaged in or exercised by nondiscursive creatures, and yet which can be algorithmically elaborated into the discursive capacity to use concepts and speak an autonomous language. But fundamental pragmatism need not take such a strong, reductive form. One might claim, more modestly, that discursive activity, from everyday thought to the cogitations of the theoretical physicist, is a species of practical intentionality (or a determination of that determinable), and indeed, one that is intelligible as having developed out of nondiscursive practical intentionality, while still maintaining that it is a wholly distinctive variety.

Fundamental pragmatism in this sense gives a distinctive shape to the naturalism of the classical American pragmatists. For that methodological commitment ensures that their naturalism is in the first instance a naturalism concerning the *subjects* of discursive understanding and agency. When we think today about naturalism, we tend to think of it first as

---

<sup>3</sup> In Chapter 3 of *Between Saying and Doing* [Oxford University Press, 2008].

a thesis about the *objects* represented by different potentially puzzling kinds of concepts: semantic, normative, probabilistic concepts, and so on. The question is how to see what those concepts represent as part of the natural world, as conceived by fundamental physics, or some special sciences, or even just by unproblematic empirical descriptive concepts. By contrast to this object naturalism, the American pragmatists were subject naturalists.<sup>4</sup> Fundamental pragmatism counsels looking first to what discursive subjects are *doing*, to the abilities they exercise, the practices they engage in. If a naturalistic story can be told about *that*, it might well be that no questions remain that should trouble the naturalist. One of the points of the toy Sprachspeile that the later Wittgenstein constructs seems to be a fundamental pragmatist, subject naturalist one—which the distinction between subject and object naturalism shows to be entirely compatible with the claim he makes already in the *Tractatus* and never relinquishes, that “philosophy is not one of the natural sciences.” Not everything we think or say need be understood as representing the world as being some way. And if it is, fundamental pragmatism invites us to understand representation in terms of what discursive subjects must *do* in order to count thereby as representing, as taking or treating some state, episode, or performance *as* a representation *of* something. For representational content is explicit—believing that things are thus-and-so. And that is to make sense of in terms of what is implicit in what the subjects do in virtue of which it is correct to say of them that they are believing that. Fundamental pragmatism is opposed to a *representationalist* order of explanation: one that begins with a notion of representational content, and appeals to that to make sense of what it is knowing and acting subjects do. That is not to say that pragmatists in this sense can have no truck at all with the concept of representation. It is to say at most that talk of representation should come at the end of the story, not the beginning.

Once a contrast between skillful *practice* and explicit *representation* has been put in place and the issue raised of their relative explanatory priority in the context of different

---

4 This is Huw Price’s terminology in “Naturalism without representationalism” in David Macarthur and Mario de Caro (eds), *Naturalism in Question* (Harvard University Press, 2004), 71–88, and (with David Macarthur) “Pragmatism, quasi-realism and the global challenge” In Cheryl Misak, ed., *The New Pragmatists* (OUP, 2007), 91–120. The other essays in his *Naturalism Without Mirrors* [Oxford University Press, 2009] can also be consulted with profit in this connection.



enterprises, the question of the relation between fundamental pragmatism and cognitive science arises. For cognitive science had as something like its original charter distinguishing its approach from that of behaviorism by its realization of the explanatory power precisely of appealing to *representations* to explain various practical cognitive abilities. Thinking about the fundamental pragmatism motivating Heidegger in setting out the project of *Being and Time*, Dreyfus drew the conclusion that the methodology of cognitive science is incompatible with the insights of that pragmatism. Is he right?

Here I think the beginning of wisdom is the realization that it makes a *big* difference whether we are talking about representations, rules, and explicitness at the *personal* level, or at the *sub*-personal level. This is in part a matter of whether one construes the *rules* the platonist invokes to articulate proprieties of practice as being *followed* by the one whose practice is in question (which would be at the personal level). Cognitive science, by contrast, postulates *sub*-personal representations, whose role is in *causal* explanations of various capacities. The sense in which they *guide* the practice is *causal*, not in the first instance *normative*. It is not at all clear that there is (or at any rate needs to be) a clash between fundamental pragmatism at the personal level and cognitive science's invocation of representations at the sub-personal level—as Dreyfus at least sometimes seems to think there is. Here one important issue is what one means by 'explicit' when fundamental pragmatism is articulated in terms of the implicit in practice vs. explicit in principle, rule, or representation form. Representations of rules are crucial for one to count as *following* a rule (as Sellars insists). In that context, representation can be thought of as the form of explicitness. But it is *not* a good idea to explicate explicitness in terms of representation if one is thinking of representation in the sense that is appropriate to the *sub*-personal level. Here the notion of specifically propositional representations is key. It is open to the pragmatist to claim (with Davidson and the author of *Making It Explicit*) that *nothing* at the sub-personal level deserves to count as *propositionally* contentful in the sense that *personal* level representations can be propositionally contentful. Belief on such a view is not a sub-personal level concept.

In order to understand the relations between fundamental pragmatism and the representational approach of cognitive science we should distinguish three levels:

a) Sub-personal representations,

- b) Practical abilities (practices) that are cognitive in some broad sense,
- c) Personal level representations.

(c) is the *explicit* properly propositional level, at which rules and principles are formulated that can express what is *implicit* at level (b). Level (b) is *practical* intentionality, and level (c) is *discursive* intentionality. Level (a) *causally* explains level (b)—and a lot of cognitive science is concerned with how this can be done in detail. The fundamental pragmatist claim is that level (c) is to be understood, explained, or explicated in terms of level (b). Cognitive science is in the business of postulating inner sub-personal representations in order to explain various kinds of skillful practice or ability. Dreyfus seems to think that approach is incompatible with the sort of fundamental pragmatism that the early Heidegger (and the later Wittgenstein) endorse. But such a view is mistaken. What that pragmatism is incompatible with is seeking to explain (b) in terms of (c), *not* (b) in terms of (a).<sup>5</sup>

#### IV. Instrumental Pragmatism

One of the variant rough formulations I offered of the methodological commitment I have called “fundamental pragmatism” is to think about norms explicitly represented in the form of rules or principles only in the context of a prior understanding of norms implicit in practice. This characterization has the advantage of placing fundamental pragmatism in the context of the Kantian normative turn, as I have claimed it should be when we think about the classical American pragmatists. The master argument for fundamental pragmatism about the normative dimension of intentionality is a regress argument familiar from the later Wittgenstein. In a nutshell, it is that the very idea of norms explicitly represented as rules or principles presupposes that of norms implicit in practices. For applying a rule is itself something that can be done

---

<sup>5</sup> In *Between Saying and Doing* [Oxford University Press, 2008] I explore the significance of the choice of the vocabulary used to specify the practices-or-abilities appealed to at level (b). This is, it seems to me, equally significant for the two enterprises, both the one that seeks to explain them and the one that seeks to use them to explain something else.

correctly or incorrectly. If we can only understand that normative assessment in turn as a matter of applying some other rule (what Wittgenstein calls an “interpretation” [Deutung]), then we are embarked on a fruitless regress. This, too, is a point that Kant had already appreciated, as an integral part of his ground-breaking normative construal of concepts as rules (for judging):

If understanding in general is to viewed as the faculty of rules, judgment will be the faculty of subsuming under rules; that is, of distinguishing whether something does or does not stand under a given rule (*casus datae legis*). General logic contains and can contain no rules for judgment...If it sought to give general instructions how we are to subsume under these rules, that is, to distinguish whether something does or does not come under them, that could only be by means of another rule. This in turn, for the very reason that it is a rule, again demands guidance from judgment. And thus it appears that, though understanding is capable of being instructed, and of being equipped with rules, judgment is a peculiar talent which can be practised only, and cannot be taught.<sup>6</sup>

The normative fundamental pragmatism of the classical American pragmatists joins cognitive science in rejecting the descriptive-dispositional behaviorism of Watson, Skinner, and Ryle. But it does so for different reasons: because of the failure of the latter group to appreciate the essentially normative character of the practical intentionality that forms the background of discursive intentionality, rather than because of their hostility to the postulation of inner representations. Wittgenstein has been called a ‘behaviorist’, in part because of his antipathy towards some kinds of explanations that appeal to inner representations. A principal danger of talking this way is that it invites overlooking his emphasis not only on the *social*, but especially on the *normative* character of the practical intentionality in the context of which he urges us to think about discursive intentionality. In this regard, Wittgenstein belongs in a box with the classical American pragmatists, and with Kant, not with the reductive descriptive-dispositional behaviorists.

---

<sup>6</sup> *Critique of Pure Reason*, A132/ B171. I discuss this regress argument further in Chapter One of *Making It Explicit* [Harvard University Press, 1994], in the context of arguments against the twin dangers of regulism and regularism about discursive norms.

But how, exactly, do the classical American pragmatists understand the basic kind of normativity implicit in practical intentionality: the kind of skillful know-how, as a species of which we are to understand discursive intentionality and its distinctive kind of normativity? I think it is not so easy to extract a clear answer to this question, even from Dewey, who has the most sophisticated approach to it. It is clear that in the most general terms the response takes the form of an appeal to the selectional-adaptional structure common to learning and evolution. The norms characteristic of the kind of practical intentionality in terms of which we are to understand discursive intentionality are immanent to and elaborated within the development of courses of experience that display this structure. In our own time, we have examples of how to make an account along these lines work. A splendid instance is Ruth Millikan's sophisticated and nuanced construction of norms in the form of Proper Functions, defined by modal counterfactual claims about selectional processes shaping reproductive families of traits.<sup>7</sup> (Millikan, a Sellars student, self-consciously takes her inspiration from Charles Morris, to whom her book is dedicated. Morris was a student of George Herbert Meade, who was in turn a student of James and a colleague of Dewey's.) I think there is every reason to believe that all of the classical American pragmatists (as well as the successors just mentioned) would have welcomed and embraced her careful working-out of their underlying idea. But of course, that detailed account was not available to them. In its absence, they often enough fall into formulations that have, from the very beginnings of the movement, led critics to attribute to the pragmatists commitment to quite a different, though not wholly unrelated, theory and to take it as the very core of the pragmatist approach.

I have in mind what is expressed by F.C.S. Schiller's slogan "The truth is what works." This is what Dewey calls "the instrumental theory" or "instrumentalism." He endorses it in such passages as these:

---

<sup>7</sup> *Language, Thought, and Other Biological Categories* [MIT Press, 1984]. The basic connection between selectional processes and alethic modal counterfactuals is indicated already by Elliot Sober's distinction between traits that are selected vs. traits that are selected *for* (*The Nature of Selection* [MIT, Bradford Press, 1984]). Millikan takes the thought much farther.

What should it mean upon the instrumental theory to accept some view or idea as true upon social credit? Clearly that such an acceptance itself works.<sup>8</sup>

What the experimentalist means is that the effective working of an idea and its truth are one and the same thing—this working being neither the cause nor the evidence of truth but its nature...<sup>9</sup>

Naturally, the pragmatist claims his theory to be true in the pragmatic sense of truth: it works, it clears up difficulties, removes obscurities, puts individuals into more experimental, less dogmatic, and less arbitrarily sceptical relations to life; aligns philosophic with scientific method; does away with self-made problems of epistemology; clarifies and reorganizes logical theory, etc. He is quite content to have the truth of his theory consist in its working in these various ways, and to leave to the intellectualist the proud possession of a static, unanalyzable, unverifiable, unworking property.<sup>10</sup>

James says such things as:

We here assume Japan to exist without ever having been there, because it WORKS to do so, everything we know conspiring with the belief, and nothing interfering...<sup>11</sup>

On pragmatistic principles, if the hypothesis of God works satisfactorily in the widest sense of the word, it is true.<sup>12</sup>

Semantic norms are understood in instrumental terms, in terms of utility. Truth-evaluable states such as beliefs are thought of on the model of tools, which can be more or less apt or useful, in concert with others that are available in a concrete situation, relative to some desired end or purpose. Taking my cue from Dewey's terminology, I'll call this approach "instrumental pragmatism" about semantic norms. There are two principal points about which it is important

---

8 "A Reply to Professor Royce's Critique of Instrumentalism" *Middle Works* Vol. 7, p. 75.

9 "The Intellectualist Criterion of Truth" *Middle Works* Vol. 4 p. 69.

10 "A Short Catechism Concerning Truth" *Middle Works* Vol 6 p. 10.

11 *Pragmatism* Lecture VI.

12 *Pragmatism* Lecture VIII.

to be clear in thinking about the instrumental strain in classical American pragmatism. First, it should be understood as at base a theory of *meaning*, not a theory of *truth*. The pragmatists did themselves no favors by pitching it in the latter way. The general idea is the fundamental pragmatist one: that the contentfulness of intentional states such as belief should be understood in terms of the contribution they make to what the believers *do*. The new element is that the doing is thought of as purposive, as aimed at some kind of end, at the satisfaction of some desire or need. Identifying success in the doing with the *truth* of the items to be thought of as contentful in virtue of their role in that process is a further, optional move. It threatens to overshadow the underlying account of meaning and content.

The second point is that that theory of the contentfulness of intentional states is a *functionalist* account. Instrumental pragmatism is a comprehensive holist functionalism about the content of states such as beliefs, desires, and intentions. It is comprehensive in that the functional systems considered comprise the organism and its whole environment. The role in such a functional system that determines the contents of states and performances caught up in it is a role in the process by which the system develops, through cycles of perception, thought, intervention transforming the environment, and perception of the results of that transaction. This is role in a course of *experience*, in what is very much a naturalized version of Hegel's sense of that term—a notion of experience that was in turn already a somewhat naturalized descendant of Kant's process of synthesis (by rational amplification, criticism, and justification) of something that exhibits the structure and unity of apperception. Processes of this sort involve felt dissatisfactions with the situation as it is at one moment, attempts to diagnose the nature of those felt dissatisfactions and to address and remove them, a process that, when all goes well, is at once the clarification of the dissatisfaction and its dissolution—the transformation of the old situation into a new one that is dissatisfying in some other way. That Kantian ancestry is particularly evident in some formulations of instrumental pragmatism. Here is one by James:

A new opinion counts as 'true' just in proportion as it gratifies the individual's desire to assimilate the novel in his experience to his beliefs in stock. It must both lean on old truth and grasp new fact; and its success... in doing this, is a matter for the individual's appreciation. When old truth grows, then, by new truth's addition, it is for subjective reasons. We are in the

process and obey the reasons. That new idea is truest which performs most felicitously its function of satisfying our double urgency. It makes itself true, gets itself classed as true, by the way it works; grafting itself then upon the ancient body of truth, which thus grows much as a tree grows by the activity of a new layer of cambium.<sup>13</sup>

Friction with stubborn reality is an integral component in this sort of process. That is the objective element in James's "double urgency." Feedback-regulated practices are 'thick', in the sense of essentially involving objects, events, and worldly states of affairs. Bits of the world are *incorporated* in such practices, in the exercise of such abilities. In this regard they contrast with words and sentences, considered merely as sign-designs or items in the natural world, which are 'thin' in that they can be specified independently of a specification of the objects or states of affairs they refer to or represent. For you cannot say what, for instance, the practice of attaching two boards with a hammer and nails is without referring to the boards, nails, and hammer. Dewey thinks of the thickness of pragmatist semantics as one of its cardinal advantages over its more traditional thin rivals. If one focuses on success as the measure of truth, rather than on functionalism about meaning, and further fails to appreciate that the functional system being considered is capacious enough to include the environment being acted on and in as well as the organism transacting with it, one will misunderstand instrumental pragmatism as a radically subjectivist view, according to which all that matters for truth is subjective feelings, and objective constraint vanishes. This is what I call "vulgar" pragmatism. James complains about this flat-footed, reductive reading already in *Pragmatism*:

Schiller says the true is that which 'works.' Thereupon he is treated as one who limits verification to the lowest material utilities. Dewey says truth is what gives 'satisfaction.' He is treated as one who believes in calling everything true which, if it were true, would be pleasant.<sup>14</sup>

---

13 *Pragmatism* Lecture II.

14 *Pragmatism* Lecture VI

And he spends most of *The Meaning of Truth* (a book that on my interpretation would better have been called *The Truth About Meaning*) rebutting that reading. Here is Dewey responding to this subjectivizing reading:

Pupil: Objection Nine. Still the pragmatic criterion, being satisfactory working, is purely personal and subjective. Whatever works so as to please me is true. Either this is your result (in which case your reference to social relations only denotes at bottom a number of purely subjectivistic satisfactions) or else you unconsciously assume an intellectual department of our nature that has to be satisfied; and whose satisfaction is truth. Thereby you admit the intellectualistic criterion.

Teacher: Reply. We seem to have got back to our starting-point, the nature of satisfaction. The intellectualist seems to think that because the pragmatist insists upon the factor of human want, purpose, and realization in the making and testing of judgments, the impersonal factor is therefore denied. But what the pragmatist does is to insist that the human factor must work itself out in cooperation with the environmental factor, and that their coadaptation is both "correspondence" and "satisfaction." As long as the human factor is ignored and denied, or is regarded as merely psychological (whatever, once more, that means), this human factor will assert itself in irresponsible ways. So long as, particularly in philosophy, a flagrantly unchastened pragmatism reigns, we shall find, as at present, the most ambitious intellectualistic systems accepted simply because of the personal comfort they yield those who contrive and accept them. Once recognize the human factor, and pragmatism is at hand to insist that the believer must accept the full consequences of his beliefs, and that his beliefs must be tried out, through acting upon them, to discover what is their meaning or consequence.<sup>15</sup>

The possible misunderstanding is, I think, actual in the reference to "our needs" as a criterion of the correctness of truth of an idea or plan. According to the essays, it

---

15 "A Short Catechism Concerning Truth" *Middle Works* Vol. 6 p. 11.



is the needs of a situation which are determinative. They evoke thought and the need of knowing, and it is only within the situation that the identification of the needs with a self occurs; and it is only by reflection upon the place of the agent in the encompassing situation that the nature of his needs can be determined. In fact, the actual occurrence of a disturbed, incomplete, and needy situation indicates that my present need is precisely to investigate, to explore, to hunt, to pull apart things now tied together, to project, to plan, to invent, and then to test the outcome by seeing how it works as a method of dealing with hard facts. One source of the demand, in short, for reference to experience as the encompassing universe of discourse is to keep us from taking such terms as "self," "my," "need," "satisfaction," etc., as terms whose meanings can be accepted and proved either by themselves or by even the most extensive dialectic reference to other terms.<sup>16</sup>

Here Dewey emphasizes not only the importance of the functionalism being comprehensive in considering a developing functional system that encompasses environment as well as the striving knower-agent, but also the holism about content that such a functionalism entails. There is no antecedently specifiable determinate content that a belief has, apart from its fellows and in advance of participating in a cycle of experience, which can then be judged *true* by pragmatist standards should the cycle conclude successfully. Rather, the belief is intelligible as having the content it does only insofar as it acquires that content by playing the role it does, along with its concomitant states, in the transactions between the believer and her world.

What goes for beliefs goes also for desires. Ends and purposes themselves are to be understood as having their content as a matter of their role in this overall system and its developmental processes. What might start out as a vague dissatisfaction itself can be clarified during the course of experience in which finding out how things are and finding out what one wants are two aspects of one process. The satisfaction of needs and wants, the achievement of goals and purposes, is the source of normativity on the instrumental construal; doing that is “working”. But what they are is (like the contents of the beliefs we are working with) itself part

---

16 Introduction to *Essays in Experimental Logic*, *Middle Works* Vol. 10, p. 364.

of what is to be determined in the course of inquiry—‘determined’ both in the sense of being made more definite and in the sense of being discovered. The former shows up from a prospective perspective, and the latter from a retrospective perspective. Dewey expended a great deal of effort in the dual process of trying to make clear and get clear himself about how the norms and standards and what they are norms and standards for assessing jointly develop in the course of experience. I cannot say that it seems to me that he succeeded very well at either task. But I do think that there is an important thought that he was after: an essentially historical perspectival structure of discursive normativity articulating a conception of determinate conceptual content that I see as also the key to understanding Hegel’s conception of experience. I have myself expended considerable effort in the dual task of trying to make that conception clear and get clear about it myself—with what success remains to be seen. I am not going to rehearse those efforts here.<sup>17</sup>

Dewey’s and James’s instrumentalism arises as one (optional) way of elaborating what is often called “Peirce’s Principle”: the meaning of a claim is the difference that adopting it would make to what one does. In fact, as I argue in Chapter Eight of *Making It Explicit*, one can get a lot more from this principle if one bifurcates it by keeping separate sets of books on the difference it makes to what one *tries* to do and difference it makes to what one *succeeds* in doing. The first of these gives one a practical difference *de dicto*. The second gives one a practical difference *de re*. Further articulating Peirce’s Principle to as to take account of the intimate *social* perspectival relations between these two sorts of practical consequence—the sense in which they are two sides of one coin—allows a much more fine-grained account of conceptual content than the classical American pragmatists managed to formulate. But that, too, is a story for another occasion. The combination of the distinction of *historical* points of view between prospective (determining as clarifying) and retrospective (determining as discovering) perspectives, and *social* points of view between attributing (*de dicto*) and acknowledging (*de re*)

---

<sup>17</sup> I talk about this structure in Chapter Three of *Reason in Philosophy* [Harvard University Press, 2009], and in Chapter Seven of *Tales of the Mighty Dead* [Harvard University Press, 2002]. It is the principal topic of my big work-in-progress on Hegel, *A Spirit of Trust*.

commitments is one of my principal suggestions for how to move forward with the ideas of the classical pragmatists.

## V. The Linguistic Turn

When classical American pragmatism is looked back upon from the perspective of the analytic movement that dominated Anglophone philosophy for at least the last half of the twentieth century, it can easily appear that a decisive wrong turn was taken after Peirce. The pragmatist founder-member was principally concerned to advance the philosophical understanding of modern logic, symbolic and natural languages, and the natural sciences—a constellation of topics that remained at the center of the analytic tradition. In his logic of relations Peirce independently achieved the bonanza of expressive power that Russell saw in Frege’s logic. But what did his successor pragmatists make of that achievement? Particularly in contrast to what Russell made of Frege, it would seem from a later vantage point that an opportunity was missed. James had little interest in logic and wrote almost nothing about it—in striking contrast to his Hegelian colleague Josiah Royce, who saw in the algebraic constructions of Alfred Bray Kempe (whom he had learned about from Peirce) a tool with which he hoped to solve the riddle of how to elaborate spatio-temporal relations from a purely conceptual basis.<sup>18</sup> The logic Dewey wrote his late, important book about was unrecognizable as such to those of his readers in 1938 whose paradigm of logic was to be found in the works of Frege, Russell, and

---

<sup>18</sup> See Bruce Kuklick’s discussion of this fascinating late project in his *Josiah Royce: An Intellectual Biography* [Hackett, 1985].

Carnap. The only pragmatist whose concern with logic matched and was recognizable as continuing that tradition was the homegrown neo-Kantian C. I. Lewis, the founder of twentieth century modal logic, who saw his own work as an attempt to synthesize the approaches of his teachers James and Royce, and in turn passed on pragmatist ideas to his students, Quine and Goodman.

Again, although James was surely the by far the best writer among the classical triumvirate, his philosophical interests focused on experience, rather than language. Dewey did write a lot about language—what he called the “tool of tools.”<sup>19</sup> He has many good things to say about the relations between meaning and use (particularly in Chapter 5 of *Experience and Nature*). But he, too, would not be recognizable to later philosophers of language as one of their number. As for science, it is not the case that James and Dewey did not care about science and the philosophy of science. But where Peirce focused on the natural sciences, James’s contributions lay on the side of psychology, and Dewey’s main interests were in the social sciences.

By “the linguistic turn” here I mean putting language at the center of philosophical concerns, and understanding philosophical problems to begin with in terms of the language one uses in formulating them. But there is a more specific significance one can take language to have. By ‘lingualism’ (compare: ‘rationalism’)—admittedly an unlovely term—is shall mean commitment to understanding *conceptual* capacities (discursiveness in general) in terms of *linguistic* capacities. Dummett epitomizes a strong version of this order of explanation:

We have opposed throughout the view of assertion as the expression of an interior act of judgment; judgment, rather, is the interiorization of the external act of assertion.<sup>20</sup>

---

<sup>19</sup> *Experience and Nature, Later Works* Vol. 1, p. 134.

<sup>20</sup> *Frege’s Philosophy of Language* [ref.] p. 361.

A weaker version of lingualism claims only that language is a *necessary* condition of discursiveness, not that it is a sufficient condition that can at least in principle be made intelligible independently of talk about discursive commitments.

It would be a mistake to conclude that the pragmatists after Peirce missed the linguistic turn. In fact, Dewey at least is clearly a (weak) lingualist about the discursive. What the pragmatists did was develop these thoughts within the context of a different approach to understanding the crucial phenomenon of language—one that was complementary to that of the analytic tradition. The Frege-Russell-Carnap approach to language takes as its paradigm artificial, formal, logistical languages articulated by explicit rules. The American pragmatists, like their fellow fundamental pragmatists the Heidegger of *Being and Time* and the Wittgenstein of the *Philosophical Investigations*, address natural languages, which they think of anthropologically, as aspects of the natural history of a certain kind of being. Their focus to begin with is not on *meaning*, but on *use*: on discursive practices, skills, and abilities, on what one must be able to *do* in order to count as saying or thinking *that* things are thus-and-so.

We can think of these two approaches as distinguished by their preferred order of explanation. The question is: which comes first, semantics (the theory of meaning) or pragmatics (the theory of use)? The logistical tradition begins with semantics: stipulating the association of some kind of semantic interpretants (paradigmatically, extensions) with basic expressions and deriving associations for more complex ones, or stipulating basic rules of derivation and then seeing what consequence relation they jointly determine. The question of how it is appropriate to *use* expressions governed by those rules is then deferred to a subsequent pragmatic theory, to which this current of thought has not traditionally devoted a great deal of attention. By contrast, the pragmatist tradition begins with pragmatics: an account precisely of how it is appropriate to use expressions. It is true that the pragmatists, also have not traditionally given a lot of attention to the specifics of the semantics that goes with such a pragmatics.

But I think we can see two principles that govern fundamental pragmatists' understanding of the relation between pragmatics and semantics. They express complementary aspects of the sense of the pragmatism in the philosophy of language that consists in insisting that semantics must answer to pragmatics. First is what I shall call "methodological pragmatism." This is the principle that the *point* of associating meanings, extensions, contents, or other semantic interpretants with linguistic expressions is to codify (express explicitly) proprieties of *use*. I think we can discern commitment to this methodological principle even in a semantic nihilist such as the later Wittgenstein. For one thing he means by saying that language is a motley is that so many and so various are the uses of any expression that there are no realistic prospects of systematizing them by associating some underlying meaning, on the basis of which one hopes then uniformly to derive the various uses (say, by one rule for declarative uses, and another for imperative ones, another for hypothetical, and so on). If the variety of uses is open-ended and unsurveyable, then there is no prospect for semantic theorizing in philosophy, precisely because the only point of such theorizing would be systematizing those proprieties of use.

The second principle governing the pragmatists' understanding of the sense in which semantics should answer to pragmatics is what I shall call "semantic pragmatism." This is the principle that in a natural language, all there is to effect the association of meanings, contents, extensions, rules, or other semantic interpretants with linguistic expressions is the way those expressions are *used* by the linguistic practitioners themselves. *Formal* semantics for *artificial* languages can content itself with the explicit stipulation of such rules or associations of meanings, by the semantic theorist working in a semantic metalanguage. *Philosophical* semantics for *natural* languages is obliged to say what it is about the practices the users of those expressions engage in or the abilities they exercise, in virtue of which they should be understood as *governed* by those rules, or as *conferring* those meanings. Semantic pragmatism is a kind of use-functionalism about meaning (the classical American pragmatists being *comprehensive* functionalists, in the sense I have given that qualification). Again, given his practice, I think commitment to such a principle can be attributed even to such a semantic pessimist as the later Wittgenstein, precisely in virtue of his criticism of various traditional ways of thinking about

meaning or content for their failure to live up to this requirement. And that sort of strategy is equally evident in Dewey's criticisms of traditional intellectualist and mentalistic conceptions.

The combination of methodological and semantic pragmatism, the two senses in which semantics can be taken to answer to pragmatics, broadly construed, might be called "linguistic pragmatism." It is one natural way of applying fundamental pragmatism to systematic theorizing about language. One of the clearest and most emphatic proponents of that conjunctive doctrine among recent philosophers is Dummett—though of course he does not associate it with pragmatism.

Quine carries forward this general pragmatist tradition in the philosophy of language when he criticizes Carnap's two-stage picture of language, according to which first meanings are stipulated, and only subsequently are theories formulated to determine which of the sentences with those meanings are true. That division of labor makes sense for artificial languages. But to understand natural languages we have to understand how the *one* thing we do, *use* the language, can serve at once to settle the meanings of our expressions *and* determine which of them we take to be true. Linguistic practice is not illuminated by postulating language/theory or meaning/belief distinctions of the Carnapian kind. As Quine famously concludes an early essay on Carnap:

"The lore of our fathers is a fabric of sentences...It is a pale grey lore, black with fact and white with convention. But I have found no substantial reasons for concluding that there are any quite black threads in it, or any white ones."<sup>21</sup>

In fact, though he did not know it, in making this pragmatist point against Carnap, Quine was recapitulating one of the important ways in which Hegel moves beyond Kant. For Kant, all our empirical activity, cognitive and practical, is discursive activity. In endorsing judgeable contents and practical maxims, knowers and agents are applying concepts. Though further

---

21 "Carnap on Logical Truth", p. 406 [ref.]

concepts may be developed thereby, for instance by judgments of reflection, one must always already have concepts in order to be apperceptively *aware* of anything at all. Hegel thought Kant was uncharacteristically, but culpably, uncritical about the origins of our primordial concepts. The locus of those concepts, Hegel thought, lies in *language*, not in some kind of experience understood as prelinguistic. Language, he said, is the existence [Dasein] of Geist—that is, of the whole normatively articulated discursive realm.<sup>22</sup> Compare Dewey:

Language in its widest sense—that is, including all means of communication such as, for example, monuments, rituals, and formalized arts—is the medium in which culture exists and through which it is transmitted.<sup>23</sup>

For Hegel, no less than for Quine and Dewey, we must understand linguistic practices as *both* instituting conceptual norms *and* applying them.<sup>24</sup> It is precisely by applying concepts in judging and acting that conceptual content is both made more determinate, going forward, and shows up as always already determinate (in the only sense in which conceptual contents are determinate), looking back.<sup>25</sup>

## VI. Rationalism and Pragmatism

---

<sup>22</sup> *Phenomenology of Spirit* [652], [666].

<sup>23</sup> *Logic, the Theory of Inquiry, Later Works* Vol. 12, p. 28

<sup>24</sup> Here are some characteristic passages:

It is therefore through culture that the individual acquires standing and actuality. His true *original nature* and substance is the alienation of himself as Spirit from his *natural* being. This individuality *moulds* itself by culture into what it intrinsically is. [I: 489]

What, in relation to the single individual, appears as his culture, is the essential moment of the substance itself, viz. the immediate passage of the [mere] thought-form of its universality into actuality; or, culture is the simple soul of the substance by means of which, what is implicit in the substance, acquires an acknowledged, real existence. The process in which the individuality moulds itself by culture is, therefore, at the same time the development of it as the universal, objective essence, i.e. the development of the actual world. Although this world has come into being through individuality, it is for self-consciousness immediately an alienated world which has the form of a fixed and solid reality over against it. [PG 490]

<sup>25</sup> See footnote 16.



Pragmatists who have made the linguistic turn take it that the most important feature of the natural history of creatures like us is that we have *come into language*<sup>26</sup>: come to engage in distinctively linguistic practices and to exercise distinctively linguistic abilities. This is both an ontogenetic and a phylogenetic achievement. Understanding it requires, at a minimum, addressing three large, interconnected kinds of question. These concern the issues of *demarcation, emergence, and leverage*. The demarcation question is definitional. How are *linguistic* practices and abilities (and hence, the lingualist about discursivity claims, *discursive* ones) to be distinguished from nonlinguistic ones? The emergence question concerns the requirement that any account of language that aspires to being naturalistic in even a very broad sense must explain the possibility of the transition from nonlinguistic to linguistic practices and abilities. How are the abilities we can see in non- or prelinguistic creatures recruited, deployed, and transformed so as to amount to linguistic ones? The leverage question is how to characterize and explain the massive qualitative difference in capacity between linguistic and nonlinguistic creatures: the bonanza of new abilities and possibilities that language opens up for those that do make the transition.

One of the principal accomplishments of the classical American pragmatists is the attention they gave to the problem of emergence, to displaying the continuities that make it naturalistically intelligible that species and individuals should be able to cross the boundary separating the prelinguistic from the linguistic. In *Experience and Nature*, Dewey sets the emergence problem this way:

Upon the whole, professed transcendentalists have been more aware than have professed empiricists of the fact that language makes the difference between brute and man. The trouble is that they have lacked naturalistic conception of its origin and status.<sup>27</sup>

In his *Logic*, he expands on this thought:

---

<sup>26</sup> We have come to see that there are substantial, potentially controversial presuppositions involved in characterizing this in terms of language *learning*.

<sup>27</sup> *Experience and Nature, Later Works* Vol. 1, p. 134.

Any theory that rests upon a naturalistic postulate must face the problem of the extraordinary differences that mark off the activities and achievements of human beings from those of other biological forms. It is these differences that have led to the idea that man is completely separated from other animals by properties that come from a non-natural source....The development of language (in its widest sense) out of prior biological activities is, in its connection with wider cultural forces, the key to this transformation. The problem, so viewed, is not the problem of the transition of organic behavior into something wholly discontinuous with it—as is the case when, for example, Reason, Intuition and the A priori are appealed to for explanation of the difference. It is a special form of the general problem of continuity of change and the emergence of new modes of activity—the problem of development at any level.<sup>28</sup>

The hallmark of an untenable intellectualism, he thinks, is an appeal to an inexplicable saltation: the ultimately miraculous dawning of consciousness or self-consciousness, the infusion of reason into a brute. The desire to provide a more satisfactory response to the emergence question than that sort of cartesian approach can offer binds Dewey together with the later Wittgenstein in a common enterprise. The point of many of the toy Sprachspiele the latter describes is to show us how features of discourse that might seem mysterious in a sense that calls for the invocation of a cartesian discontinuity can be exhibited already in practices we can see that intelligent nonlinguistic hominids could master.

When we turn to the demarcation question, however, I think the pragmatists disappoint. What is distinctive of linguistic (or discursive) practices? What sets them apart from prelinguistic or nondiscursive practices? It is one's answer to this question that ties together the

---

<sup>28</sup> *Logic, the Theory of Inquiry, Later Works* Vol. 12, p. 50. This emphasis on continuity does not lead Dewey to ignore the differences that language makes:

The evidence usually adduced in support of the proposition that lower animals, animals without language, think, turns out, when examined, to be evidence that when men, organisms with power of social discourse, think, they do so with the organs of adaptation used by lower animals, and thus largely repeat in imagination schemes of overt animal action. But to argue from this fact to the conclusion that animals think is like concluding that because every tool, say a plow, originated from some pre-existing natural production, say a crooked root or forked branch, the latter was inherently and antecedently engaged in plowing. The connection is there, but it is the other way around. *Experience and Nature, Later Works* Vol. 1, p. 215.

emergence question with the leverage question. For the criteria of adequacy for answers to those questions turn on its being the *same* kind of practices and abilities that one has told a story about the nonmiraculous emergence of, in answering the first question, that one then must show can intelligibly account for the huge differences in capabilities, cognitive and practical, that come with the advent of language, in answering the second question. We need not assume that the emergence of language is an all-or-none thing. One might, with Wittgenstein, want to deny that there is or need be a bright line separating the discursive from the nondiscursive, in favor of a family-resemblances sort of view. A pluralist-incrementalist response to the demarcation question makes the emergence question easier to answer, but makes the leverage question correspondingly more difficult. I don't think Dewey's metainstrumentalist "tool of tools" line can be made to work to bring the emergence and leverage issues into harmony—but I've argued that elsewhere and won't rehearse my complaints here.<sup>29</sup> Apart from that, he seems to offer only vague remarks about language as a making enhanced the possibilities of co-operation and rising above the individual standpoint.<sup>30</sup>

I cannot here address the all-important leverage question.<sup>31</sup> But the demarcation question is prior. After all, if one is going to say how Geist precipitates out of nature, and how it transforms sentient organisms into sapient ones, one should try to say what it is. The challenge is to offer satisfactory responses to *both* the emergence question *and* the leverage question. Focusing on just one of them makes it too easy. In the passage above, Dewey says in effect that the neo-cartesian intellectualists make the leverage question too easy to respond to, by ignoring (or making it impossible to address) the question of emergence. I have just accused him of

---

<sup>29</sup> *Experience and Nature, Later Works*, Vol. 1, p. 134. I discuss this approach in Chapters One and Two.

<sup>30</sup> I have in mind passages such as this one:

The importance of language as the necessary, and, in the end, sufficient condition of the existence and transmission of non- purely organic activities and their consequences lies in the fact that, on one side, it is a strictly biological mode of behavior, emerging in natural continuity from earlier organic activities, while, on the other hand, it compels one individual to take the standpoint of other individuals and to see and inquire from a standpoint that is not strictly personal but is common to them as participants or "parties" in a conjoint undertaking. It may be directed by and towards some physical existence. But it first has reference to some other person or persons with whom it institutes communication--the making of something common. Hence, to that extent its reference becomes general and "objective."<sup>30</sup> *Logic, the Theory of Inquiry, Later Works* Vol. 12, p. 52.

<sup>31</sup> I do address it in *Making It Explicit* and *Between Saying and Doing*.

making the complementary mistake. In any case, it is clear that the hinge that connects the issues of emergence and leverage is the question of demarcation. For the challenge is to show that the *same* phenomenon that one has accounted for the emergence of can leverage sentience into sapience. So demarcating the realm of linguistic or discursive practices and abilities is an absolutely essential element of the philosophical project I have been describing.

I want to close with a suggestion as to one way fundamental pragmatists (those committed to understanding discursive intentionality as a kind of practical intentionality) who are weak lingualists about discursiveness (take engaging in linguistic practices as a necessary condition of deploying concepts)—a class I take to include at least Peirce, Dewey, the early Heidegger, and the later Wittgenstein—might answer the demarcation question, and so determine definite criteria of adequacy for responses to the leverage question. My idea is that pragmatism can usefully be combined with a *rationalist* criterion of demarcation of the linguistic—and hence of discursiveness in general. By this I mean that what distinguishes the linguistic practice in virtue of which we are sapient and not merely sentient beings is its core practices of giving and asking for *reasons*. A necessary and sufficient condition of being a discursive practice is that some performances are accorded by it the pragmatic significance of *claimings* or assertings. Semantically, claimable or assertible contents are *propositional* contents. Syntactically, what expresses those contents is declarative sentences. This combination of pragmatic, semantic, and syntactic features is the iron triangle of discursiveness. The pragmatist order of explanation of course starts with the pragmatics. The thought is that to have the pragmatic significance of an assertion is to be able both to serve as a reason, and potentially to stand in need of reasons. So propositional contents are those that can play the role both of premise and of conclusion in *inferences*. Discursive practice is accordingly understood as essentially inferentially articulated.

In *Logic: The Theory of Inquiry*, the normative status with which Dewey's pragmatics begins, in terms of which the semantics is to be articulated, is *assertibility*. I have argued on the one hand that to be recognizable as engaging in a practice of making claims and (so) giving and asking for reasons, a community must distinguish at least *two* normative statuses: commitment

and entitlement to commitments, and further, that splitting the single status of assertibility into these two aspects pays huge benefits semantically.<sup>32</sup> Specifically, one can use them to define three kinds of material inference: commitment-preserving inferences, entitlement-preserving inferences, and incompatibility entailments. The core of my strong inferentialist version of rationalistic pragmatism lies in the claim that conceptual content consists in inferential role in a broad sense, articulated along those three dimensions.<sup>33</sup> Of course the underlying rationalist criterion of demarcation of the discursive could be worked out in other ways.

Commitment to a rationalist criterion of demarcation of the discursive requires disagreeing with Wittgenstein: Language *does* have a downtown, and it is the practice of making claims and giving and asking for reasons. Other things we can do with language are ancillary to and parasitic upon these essential core functions. On this view, most of the toy practices Wittgenstein calls “Sprachspiele” are *vocal*, but not genuinely *verbal*, not really *language* games. The builder’s utterances in the opening ‘Slab’ practice, for instance, should not be understood as imperatives. They are vocalizations that have the pragmatic significance of making certain responses on the part of the assistant appropriate. But genuine imperatives do that by *saying* what it is that ought to be done. In this full-blooded sense, no practice can contain the genuine imperative “Bring be a slab,” unless it also contains declaratives such as “This is a slab.”

Wittgenstein and Dewey are together in rejecting rationalist criteria of demarcation of the linguistic (and hence the discursive)—indeed, in resisting offering any answer at all to the demarcation question. In Dewey’s case, the idea of a rationalist pragmatism would probably have struck him as a *contradictio in adjecto*. But rationalism as I have described it is not a form of the intellectualism that stands opposed to fundamental pragmatism. It is wholly compatible with understanding discursive intentionality as a kind of practical intentionality: specifically, as the kind that includes practices of making claims and giving and asking for reasons. It aims to

---

<sup>32</sup> See Chapter Six of my *Articulating Reasons* [Harvard University Press, 2001]

<sup>33</sup> For the distinction between weak, strong, and hyperinferentialism, see the Introduction to *Articulating Reasons*. Inferentialism is just one form that rationalism might take. For there is more to reason than inference. Making distinctions, formulating definitions, and producing constructions are all rational processes, alongside drawing conclusions.

say what structure a norm-instituting social practice must have in order properly to be understood as such a practice. It offers a specific proposal for how to understand the kind of practical knowing *how* that adds up to cognitive claiming *that*: it is practical mastery of inferential relations and transitions. And answering the demarcation question about discursive practice in a rationalist manner neither makes it impossible in principle to answer the emergence question nor obliges one to give a cartesian answer to it.

I began my story about pragmatism in an unconventional place: with Kant's *normative* criterion of demarcation of the discursive, that is, with his idea that what is distinctive of judgments and intentional actions is that they are things we are *responsible* for. They are kinds of *commitments*. But that normative criterion of demarcation was also a *rationalist* criterion of demarcation. For he understood that responsibility, that commitment, as a *rational* responsibility, as the *justificatory* responsibility to have reasons for ones theoretical and practical commitments, the *ampliative* responsibility to acknowledge their inferential consequences, and the *critical* responsibility to revise commitments that are incompatible, that is, that serve as reasons against one another. Kant's pragmatism consists in his strategy of understanding semantic content in terms of what apperceiving subjects must *do* to fulfill those responsibilities. Judgeable contents have to stand to one another in relations of material consequence and incompatibility: the inferential relations that constrain the process of synthesizing a constellation of commitments and entitlements exhibiting the distinctive unity of apperception. Wittgenstein's example teaches that we should follow Hegel's steps toward naturalizing Kant's notion of norms by understanding norms as implicit in social practices. Normative statuses of responsibility and commitment are *social* statuses: creatures of our practical attitudes of taking or treating each other *as* responsible and committed.

The move beyond Dewey and Wittgenstein to a rationalist, more specifically inferentialist pragmatism that I am recommending is accordingly also a return to pragmatism's roots in German idealism. As Kant synthesized empiricism and rationalism, and the pragmatists synthesized naturalism and empiricism, I'm suggesting that a way forward is to synthesize pragmatism and rationalism—in the form of the rationalist response to the demarcation question.



## 2014 Nordic Pragmatism Lectures:

### Analytic Pragmatism, Expressivism, and Modality

#### Lecture 2

### On the Way to a Pragmatist Theory of the Categories

#### 1. Introduction

Several decades ago, Richard Rorty suggested that philosophical admirers of Wilfrid Sellars could be divided into two schools, defined by which of two famous passages from his masterwork “Empiricism and the Philosophy of Mind” are taken to express his most important insight. The two passages are:

In the dimension of describing and explaining the world, science is the measure of all things, of what is that it is, and of what is not that it is not" (§41).

and

[In] characterizing an episode or a state as that of *knowing*, we are not giving an empirical description of that episode or state; we are placing it in the logical space of reasons, of justifying and being able to justify what one says” (§36).<sup>34</sup>

The first passage, often called the “*scientia mensura*,” expresses a kind of scientific naturalism. Its opening qualification is important: there are other discursive and cognitive activities besides

---

<sup>34</sup> In Herbert Feigl and Michael Scriven, eds., *Minnesota Studies in the Philosophy of Science*, vol. I (Minneapolis: University of Minnesota Press, 1956); reprinted in Sellars’s *Science, Perception, and Reality* (London: Routledge and Kegan Paul, 1956; reissued Atascadero, Ridgeview, 1991); reprinted as a monograph, with an Introduction by Richard Rorty and a Study Guide by Robert Brandom (Cambridge, Mass.: Harvard University Press, 1997).



describing and explaining. The second passage says that characterizing something as a knowing is one of them. And indeed, Sellars means that in characterizing something even as a believing or a believable, as conceptually contentful at all, one is doing something other than describing it. One is placing the item in a normative space articulated by relations of what is a reason for what. Meaning, for him, is a normative phenomenon that does not fall within the descriptive realm over which natural science is authoritative.

Rorty called those impressed by the scientific naturalism epitomized in the *scientia mensura* “right wing Sellarsians,” and those impressed by the normative nonnaturalism about semantics expressed in the other passage “left wing Sellarsians.” Acknowledging the antecedents of this usage, he used to express the hope that right wing and left wing Sellarsians would be able to discuss their disagreements more amicably and irenically than did the right wing and left wing Hegelians, who, as he put it, “eventually sorted out their differences at a six-month-long seminar called ‘the Battle of Stalingrad.’” According to this botanization, I am, like my teacher Rorty and my colleague John McDowell, a left wing Sellarsian, by contrast to such eminent and admirable right wing Sellarsians as Ruth Millikan, Jay Rosenberg, and Paul Churchland.

While I think Rorty’s way of dividing things up is helpful (there really are “41-ers” and “36-ers”), I want here to explore a different perspective on some of the same issues. I, too, will focus on two big ideas that orient Sellars’s thought. I also want to say that one of them is a good idea and the other one, on the whole, a bad idea—a structure that is in common between those who would self-identify as either right- or left-wing Sellarsians. And the one I want to reject is near and dear to the heart of the right wing. But I want, first, to situate the ideas I’ll consider in the context of Sellars’s neo-Kantianism: they are his ways of working out central ideas of Kant’s. Specifically, they are what Sellars makes of two fundamental ideas that are at the center of Kant’s transcendental idealism: the metaconcept of *categories*, or *pure concepts of the understanding*, and the distinction between *phenomena* and *noumena*. The latter is a version of the distinction between appearance and reality, not in a light epistemological sense, but in the ontologically weighty sense that is given voice by the *scientia mensura*. I cannot say that these

fall under the headings, respectively, of What Is Living and What Is Dead in Sellars's thought, since the sort of scientific naturalism he uses to interpret Kant's phenomena/noumena distinction is undoubtedly very widespread and influential in contemporary Anglophone philosophy. My aim here is threefold: to explain what I take it Sellars makes of these Kantian ideas, why I think the first line of thought is more promising than the second, and the way forward from each that seems to me most worth developing.

When asked what he hoped the effect of his work might be, Sellars said he would be happy if it helped usher analytic philosophy from its Humean into its Kantian phase. (*A propos* of this remark, Rorty also said, not without justice, that in these terms my own work could be seen as an effort to help clear the way from analytic philosophy's incipient Kantian phase to an eventual Hegelian one.<sup>35</sup>) Sellars tells us that his reading of Kant lies at the center of his work. He used that theme to structure his John Locke lectures, to the point of devoting the first lecture to presenting a version of the Transcendental Aesthetic with which Kant opens the *Critique of Pure Reason*. Those lectures, published as *Science and Metaphysics: Variations on Kantian Themes*, are Sellars's only book-length, systematic exposition of his views during his crucial middle period. The development of Kantian themes is not only self-consciously used to give that book its distinctive shape, but also implicitly determines the contours of Sellars's work as a whole. I think the best way to think about Sellars's work is as a continuation of the neo-Kantian tradition. In particular, I think he is the figure we should look to today in seeking an appropriation of Kant's theoretical philosophy that might be as fruitful as the appropriation of Kant's practical philosophy that Rawls initiated. On the theoretical side, Sellars was the greatest neo-Kantian philosopher of his generation.<sup>36</sup>

In fact, the most prominent neo-Kantians of the previous generation: C. I. Lewis and Rudolf Carnap were among the most immediate influences on Sellars's thought. Kant was the

---

<sup>35</sup> In his Introduction to my Harvard University Press edition of "Empiricism and the Philosophy of Mind."

<sup>36</sup> His only rival for this accolade, I think, would be Peter Strawson, who certainly did a lot to make us realize that a reappropriation of some of Kant's theoretical philosophy might be a viable contemporary project. But I do not think of Peter Strawson's work as *systematically* neo-Kantian in the way I want to argue that Sellars's is.

door through which Lewis found philosophy, and later, the common root to which he reverted in his attempt to reconcile what seemed right to him about the apparently antithetical views of his teachers, William James and Josiah Royce. (Had he instead been trying to synthesize Royce with Dewey, instead of James, he would have fetched up at Hegel.) In his 1929 *Mind and the World Order*, Lewis introduced as a central technical conception the notion of the sensory “Given”, which Sellars would famously use (characteristically, without mentioning Lewis by name) as the paradigm of what he in *EPM* called the “Myth of the Given.” (Indeed, shortly after his 1946 *An Analysis of Knowledge and Valuation*, which Sellars also clearly has in mind in *EPM*, Lewis wrote a piece addressing the question “Is The Givenness of the Given Given?” His answer was No: It is a necessary postulate of high philosophical theory, which dictates that without an sensory Given, empirical knowledge would be impossible.)

We shall see in subsequent lectures that Sellars modeled his own Kantian “metalinguistic” treatments of modality and the ontological status of universals explicitly on ideas of Carnap. Although, like Lewis, Carnap is not explicitly mentioned in *EPM*, his presence is registered for the philosophical cognoscenti Sellars took himself to be addressing there by the use of the Carnapian term “protocol sentence” (as well as Schlick’s “Konstatierung”) for noninferential observations. Unlike Lewis, Carnap actually stood in the line of inheritance of classical nineteenth-century German neo-Kantianism. His teacher, Bruno Bauch, was (like Heidegger), a student of Heinrich Rickert in Freiburg—who, with the older Wilhelm Windelband, led the Southwest or Baden neo-Kantian school. In spite of these antecedents, Bauch was in many ways closer to the Marburg neo-Kantians, Hermann Cohen and Paul Natorp, in reading Kant as first and foremost a philosopher of the natural sciences, mathematics, and logic. I suppose that if one had asked Carnap in what way his own work could be seen as a continuation of the neo-Kantian tradition of his teacher, he would first have identified with this Marburg neo-Kantian understanding of Kant, and then pointed to the *logical* element of his logical empiricism—itsself a development of the pathbreaking work of Frege, Bauch’s friend and colleague at Jena when Carnap studied with both there—as giving a precise and modern form to the conceptual element in empirical knowledge, which deserved to be seen as a worthy successor to Kant’s own version of the conceptual.

If Lewis and Carnap do not immediately spring to mind as neo-Kantians, that is because each of them gave Kant an empiricist twist, which Sellars was concerned to undo. If you thought that Kant thought that the classical empiricists' Cartesian understanding of the sensory contribution to knowledge was pretty much all right, and just needed to be supplemented by an account of the independent contribution made by a conceptual element, you might well respond to the development of the new twentieth century logic with a version of Kant that looks like Lewis's *Mind and the World Order*, and *An Analysis of Knowledge and Valuation*, and Carnap's *Aufbau* (and for that matter, Nelson Goodman's *Structure of Appearance*). That assumption about Kant's understanding of the role played by sense experience in empirical knowledge is exactly what Sellars challenges in *EPM*.

One of the consequences of his doing that is to make visible the neo-Kantian strand in analytic philosophy that Lewis and Carnap each, in his own way, represented—and which Sellars and, in our own time, John McDowell further developed. Quine was a student of both Lewis and Carnap, and the Kantian element of the common empiricism he found congenial in their thought for him drops out entirely—even though the logic remains. His Lewis and his Carnap are much more congenial to a narrative of the history of analytic philosophy initiated by Bertrand Russell and G.E. Moore, according to which the movement is given its characteristic defining shape as a recoil from Hegel (seen through the lenses of the British Idealism of the waning years of the nineteenth century). They understood enough about the Kantian basis of Hegel's thought to know that a *holus bolus* rejection of Hegel required a diagnosis of the idealist rot as having set in already with Kant. This narrative does pick out one current in the analytic river—indeed, the one that makes necessary the reappropriation of the metaconceptual resources of Kant's theoretical philosophy in the late twentieth and early twenty-first centuries. But it was never the whole story.<sup>37</sup> The neo-Kantian tradition comprising Lewis, Carnap, and Sellars can be thought of as an undercurrent, somewhat occluded from view by the empiricist surface.

---

<sup>37</sup> Paul Redding begins the process of recovering the necessary counter-narrative in the Introduction to his *Analytic Philosophy and the Return of Hegelian Thought* [Cambridge University Press, 2010].

## 2. Categories in Kant

Many Kantian themes run through Sellars's philosophy. I am going to focus on one master-idea, each of which orients and ties together a number of otherwise apparently disparate aspects of his work. It is the idea that besides concepts whose characteristic expressive job it is to describe and explain empirical goings-on, there are concepts whose characteristic expressive job it is to make explicit necessary structural features of the discursive framework within which alone description and explanation are possible. Failing to acknowledge and appreciate this crucial difference between the expressive roles different bits of vocabulary play is a perennial source of distinctively philosophical misunderstanding. In particular, Sellars thinks, attempting to understand concepts doing the second, framework-explicating sort of work on the model of those whose proper use is in empirical description and explanation is a fount of metaphysical and semantic confusion.<sup>38</sup> Among the vocabularies that play the second sort of role, Sellars includes *modal* vocabulary (not only the alethic, but also the deontic species), *semantic* vocabulary, *intentional* vocabulary, and *ontological-categorical* vocabulary (such as 'proposition', 'property' or 'universal', and 'object' or 'particular'). It is a mistake, he thinks, to understand the use of any of these sorts of vocabulary as fact-stating in the narrow sense that assimilates it to *describing* how the world is. It is a corresponding mistake to recoil from the metaphysical peculiarity and extravagance of the kinds of facts one must postulate in order to understand statements couched in these vocabularies as fact-stating in the narrow sense (e.g. normative facts, semantic facts, conditional facts, facts about abstract universals) by denying that such statements are legitimate, or even that they can be true. (Though to say that they are true is not, for Sellars, to describe them.) Both mistakes (the dogmatic metaphysical and the skeptical), though opposed to one another, stem from the common root of the *descriptivist fallacy*. That is the failure to see

---

<sup>38</sup> Distinguishing two broadly different kinds of *use* bits of vocabulary can play does not entail that there are two corresponding kinds of *concepts*—even in the presence of the auxiliary Sellarsian hypothesis that grasp of a concept is mastery of the use of a word. Though I suppress the distinction between these two moves in these introductory formulations, it will become important later in the story.

that some perfectly legitimate concepts do not play a narrowly descriptive role, but rather a different, explicative one with respect to the practices of description and explanation. Following Carnap, Sellars instead analyses the use of all these kinds of vocabulary as, each in its own distinctive way, “covertly metalinguistic.”

In opposing a Procrustean descriptivism about the expressive roles locutions can play, Sellars makes common cause with the later Wittgenstein. For Wittgenstein, too, devotes a good deal of effort and attention to warning us of the dangers of being in thrall to (“bewitched by”) a descriptivist picture. We must not simply assume that the job of all declarative sentences is to state facts (“I am in pain,” “It is a fact that ...”), that the job of all singular terms is to pick out objects (“I think...,” “I have a pain in my foot,”), and so on. In addition to tools for attaching, detaching, and in general re-shaping material objects (hammer and nails, saws, draw-knives...) the carpenter’s tools also include plans, a foot-rule, level, pencil, and toolbelt. So, too, with discursive expressive ‘tools’. Wittgenstein’s expressive pluralism (language as a motley) certainly involves endorsement of the anti-descriptivism Sellars epitomizes by saying

[O]nce the tautology ‘The world is described by descriptive concepts’ is freed from the idea that the business of all non-logical concepts is to describe, the way is clear to an *ungrudging* recognition that many expressions which empiricists have relegated to second-class citizenship in discourse are not *inferior*, just *different*.<sup>39</sup>

But Sellars differs from Wittgenstein in characterizing at least a broad class of nondescriptive vocabularies as playing generically the *same* expressive role. They are broadly metalinguistic locutions expressing necessary features of the framework of discursive practices that make description (and—so—explanation) possible. Of this broad binary distinction of expressive roles, with ordinary empirical descriptive vocabulary on one side and a whole range of

---

<sup>39</sup> CDCM §79.

apparently disparate vocabularies going into another class as “metalinguistic”, there is, I think, no trace in Wittgenstein.<sup>40</sup>

The division of expressive roles that I am claiming for Sellars binds together modal, semantic, intentional, and ontological-categorial vocabulary in opposition to empirical-descriptive vocabularies traces back to Kant’s idea of “pure concepts of the understanding,” or categories, which play quite a different expressive role from that of ordinary empirical descriptive concepts. The expressive role of pure concepts is, roughly, to make explicit what is implicit in the use of ground-level concepts: the conditions under which alone it is possible to apply them, which is to say, use them to make judgments. Though very differently conceived, Kant’s distinction is in turn rooted in the epistemological difference Hume notices and elaborates between ordinary empirical descriptive concepts and concepts expressing lawful causal-explanatory connections between them. Hume, of course, drew skeptical conclusions from the observation that claims formulated in terms of the latter sort of concept could not be justified by the same sort of means used to justify claims formulated in terms of empirical descriptive concepts.

Kant, however, looks at Newton’s formulation of the best empirical understanding of his day and sees that the newly introduced concepts of force and mass are not intelligible apart from the *laws* that relate them. If we give up the claim that  $F$  equals  $m \cdot a$  then we do not mean force and mass, but are using some at least slightly different concepts. (Galileo’s geometrical version of the (late medieval) observable concept of acceleration is antecedently intelligible). This leads Kant to two of his deepest and most characteristic metaconceptual innovations: thinking of statements of laws formulated using alethic modal concepts as making explicit rules for reasoning with ordinary empirical descriptive concepts, and understanding the contents of such concepts as articulated by those rules of reasoning with them.

---

<sup>40</sup> The best candidate might be the discussion of “hinge propositions” in *On Certainty*. But the point there is, I think, different. In any case, Wittgenstein does not *generalize* the particular expressive role he is considering to anything like the extent I am claiming Sellars does.

This line of thought starts by revealing the semantic presuppositions of Hume's epistemological arguments. For Hume assumes that the contents of ordinary empirical descriptive concepts are intelligible antecedently to and independently of taking them to stand to one another in rule-governed inferential relations of the sort made explicit by modal concepts. Rejecting that semantic atomism then emerges as a way of denying the intelligibility of the predicament Hume professes to find himself in: understanding ordinary empirical descriptive concepts perfectly well, but getting no grip thereby on the laws expressed by subjunctively robust rules relating them. Even though Kant took it that Hume's skeptical epistemological argument rested on a semantic mistake, from his point of view Hume's investigation had uncovered a crucial *semantic* difference between the expressive roles of different kinds of concepts. Once his attention had been directed to them, he set himself the task of explaining what was special about these *nondescriptive* concepts.

Two features of Kant's account of the expressive role distinctive of the special class of concepts to which Hume had directed his attention are of particular importance for the story I am telling here. They are *categorial* concepts, and they are *pure* concepts. To say that they are 'categorial' in this context means that they make explicit aspects of the *form* of the conceptual as such. For Kant concepts are functions of judgment, that is, they are to be understood in terms of their role in judging. Categorial concepts express structural features of empirical descriptive judgments. What they make explicit is implicit in the capacity to make any judgments at all. This is what I meant when I said above that rather than describing how the world is, the expressive job of these concepts is to make explicit necessary features of the framework of discursive practices within which it is possible to describe how the world is. The paradigm here is the alethic modal concepts that articulate the subjunctively robust consequential relations among descriptive concepts.<sup>41</sup> It is those relations that make possible *explanations* of why one description applies because another does. That force *necessarily* equals the product of mass and acceleration means that one can explain the specific acceleration of a given mass by describing

---

<sup>41</sup> Note that these concepts are *not* those Kant discusses under the heading of "Modality", but rather concern the hypothetical form of judgment.



the force that was applied to it. (Of course, Kant also thinks that in articulating the structure of the judgeable as such, these concepts *thereby* articulate the structure of what is empirically *real*: the structure of *nature*, of the *objective world*. But this core thesis of his understanding of empirical realism within transcendental idealism is an optional additional claim, not entailed by the identification of a distinctive class of concepts as categories of the understanding.)

To say that these concepts are ‘pure’ is to say that they are available to concept-users (judgers = those who can understand, since for Kant the understanding is the faculty of judgment) *a priori*.<sup>42</sup> Since what they express is implicit in any and every use of concepts to make empirical judgments, there is no *particular* such concept one must have or judgment one must make in order to be able to deploy the pure concepts of the understanding. To say that judgers can grasp these pure concepts *a priori* is *not* to say that they are *immediate* in the Cartesian sense of nonrepresentational. Precisely not. The sort of self-consciousness (awareness of structural features of the discursive as such) they make possible is mediated by those pure concepts. What was right about the Cartesian idea of the immediacy of self-consciousness is rather that these mediating concepts are available to every thinker *a priori*. Their grasp does not require grasp or deployment of any *particular* ground-level empirical concepts, but is *implicit* in the grasp or deployment of *any* such concepts. The way I will eventually recommend that we think about this distinctive *a prioricity* is that in being able to deploy ordinary empirical descriptive concepts one already knows how to do everything one needs to know how to do in order to be able to deploy the concepts that play the expressive role characteristic of concepts Kant picks out as “categorical” (as well as some that he does not).

### 3. Categories in Sellars

Sellars’s development of Kant’s idea of pure concepts of the understanding is articulated by two master ideas. First, his successor metaconception comprises concepts that are in some

---

<sup>42</sup> I take it that Kant always uses “*a priori*” and “*a posteriori*” as adverbs, modifying some some verb of cognition, paradigmatically “know”.

broad sense *metalinguistic*.<sup>43</sup> In pursuing this line he follows Carnap, who besides ground-level empirical descriptive vocabulary allowed metalinguistic vocabulary as also legitimate in formal languages regimented to be perspicuous. Such metalinguistic vocabulary allows the formulation of explicit rules governing the use of descriptive locutions. Ontologically classifying terms such as ‘object’, ‘property’, and ‘proposition’ are “quasi-syntactical” metavocabulary corresponding to overtly syntactical expressions in a proper metalanguage such as ‘singular term’, ‘predicate’, and ‘declarative sentence’. They are used to formulate “L-rules”, which specify the structure of the language in which empirical descriptions are to be expressed.<sup>44</sup> Alethic modal vocabulary is used to formulate “P-rules”, which specify rules for reasoning with particular empirically contentful descriptive vocabulary. Carnap’s neo-Kantianism does not extend to embracing the metaconcept of categories, which he identifies with the excesses of transcendental idealism. But in the expressions Carnap classifies as overtly or covertly metalinguistic, Sellars sees the raw materials for a more thoroughly Kantian successor conception to the idea of pure categories of the understanding.

The second strand guiding Sellars’s reconceptualization of Kantian categories is his *semantic inferentialist* approach to understanding the contents of descriptive concepts. Sellars picks up on Kant’s rejection of the semantic atomism characteristic of both the British empiricism of Locke and Hume that Kant was reacting to and of the logical empiricism of Carnap that Sellars was reacting to.<sup>45</sup> The way he works out the anti-atomist lesson he learns from Kant is in terms of the essential contribution made to the contents of ordinary empirical descriptive concepts by the inferential connections among them appealed to in *explanations* of why some descriptions apply to something in terms of other descriptions that apply to it.

Although describing and explaining (predicting, retrodicting, understanding) are *distinguishable*, they are also, in an important sense, *inseparable*. It is only because the expressions in terms of which we describe objects, even such basic

---

<sup>43</sup> In Chapter Three I discuss the sense in which “metalinguistic” should be understood in such formulations..

<sup>44</sup> Chapter Seven discusses Sellars’s view about this kind of locution.

<sup>45</sup> “Another feature of the empiricist tradition is its ‘logical atomism,’ according to which every basic piece of empirical knowledge is logically independent of every other. Notice that this independence concerns not only *what* is known, but the *knowing* of it. The second dimension of this ‘atomism’ is of particular importance for understanding Kant’s rejection of empiricism...”[“Towards a Theory of the Categories” §16]

expressions as words for perceptible characteristics of molar objects, locate these objects in a space of implications, that they describe at all, rather than merely label. The descriptive and explanatory resources of language advance hand in hand.<sup>46</sup>

This is a rich and suggestive passage. It is worth unpacking the claims it contains. It is framed by a distinction between a weaker notion, labeling, and a stronger one, describing. By ‘labeling’ Sellars means discriminating, in the sense of responding differentially. A linguistic expression is used as a label if its *whole* use is specified by the *circumstances* under which it is applied—the *antecedents* of its application. We might distinguish between three kinds of labels, depending on how we think of these circumstances or antecedents. First, one could look at what stimuli as a matter of fact elicit or in fact have elicited the response that is being understood as the application of a label. Second, one could look *dispositionally*, at what stimuli *would* elicit the application of the label. Third, one could look at the circumstances in which the label is *appropriately* applied. What the three senses have in common is that they look only *upstream*, to the situations that have, would, or should prompt the use of the label. The first provides no constraint on future applications of the label—*que sera sera*—as familiar gerrymandering arguments about “going on in the same way” remind us. The second doesn’t fund a notion of mistaken application. However one is disposed to apply the label is proper, as arguments summarized under the heading of “disjunctivitis” make clear. Only the third, normatively richer sense in which the semantics of a label consists in its circumstances of *appropriate* application (however the proprieties involved are understood) makes intelligible a notion of mislabeling.

Sellars wants to distinguish labeling in *all* of these senses from *describing*. The idea is that since labeling of any of these sorts looks only to the *circumstances* in which the label is, would be, or should be applied, expressions used with the semantics characteristic of labels address at most one of the two fundamental aspects of the use characteristic of descriptions. The rules for the use of labels tell us something about what is (or would be or should be) in effect so described, but say nothing at all about what it is described *as*. That, Sellars thinks, depends on

---

<sup>46</sup> CDCM §108.

the *consequences* of applying one description rather than another. The semantics of genuine descriptions must look downstream, as well as upstream. It is this additional feature of their use that distinguishes descriptions from labels. Helab one might quibble verbally with Sellars's using 'label' and 'description' to describe expressions whose semantics depends on only one or on both of these dimensions of use. But it seems clear that a real semantic distinction is being marked.

Making a further move, Sellars understands those consequences of application of descriptions as essentially involving *inferential* connections to other descriptive concepts. This is what he means by saying that what distinguishes descriptions from labels is their situation in a "space of implications." We can think of these implications as specifying what other descriptions do, would, or should *follow from* the application of the initial, perhaps responsively elicited, description. As he is thinking of things, a description (correctly) applies to a range of things (for descriptive concepts used observationally, including those that are appropriately noninferentially differentially responded to by applying the concept), which are described *by* it. And it describes them *as* something from which a further set of descriptions (correctly) follows. Crucially, these further descriptions can themselves involve applications of descriptive concepts that also have *non-inferential* (observational) circumstances of application. Descriptive concepts that have *only* inferential circumstances of application he calls 'theoretical' concepts.

In the opening sentence of the passage Sellars includes *understanding* as one of the phenomena he takes to be intricately with description in the way explaining is. Understanding a descriptive concept requires being able to place it in the "space of implications," partly in virtue of which it has the content that it does. This is in general a kind of knowing *how* rather than a kind of knowing *that*: being able to distinguish in practice the circumstances and consequences of application of the concept, when it is appropriately applied and what follows from so applying it. Grasping a concept in this sense is not an all-or-none thing. The ornithologist knows her way around inferentially in the vicinity of terms such as 'icterid' and 'passerine' much better than I do. A consequence of this way of understanding understanding is that one cannot grasp one

concept without grasping many. This is Sellars's way of developing Kant's anti-atomist semantic insight.

Taking a further step (undertaking a commitment not yet obviously entailed by the ones attributed so far), Sellars also thinks that the inferences articulating the consequences of concepts used descriptively must always include *subjunctively robust* inferences. That is, the inferences making up the "space of implications" in virtue of which descriptive concepts have not only potentially atomistic circumstances of application but also non-atomistic relational consequences of application must extend to what other descriptions *would be* applicable if a given set of descriptions *were* applicable. For what Sellars means by 'explanation' is understanding the applicability of some descriptions as *explained by* the applicability of others according to just this kind of inference. This is, of course, just the sort of inferential connection that Hume's empiricist atomistic semantics for descriptive concepts, construing them as labels, could not underwrite. Sellars's conception of descriptions, as distinguished from labels, is his way of following out what he sees as Kant's anti-atomist semantic insight. *Modal* concepts make explicit these *necessary* inferential-consequential connections between descriptive concepts. They thereby perform the expressive role characteristic of Kantian categories: expressing essential features of the framework within which alone genuine description is possible.

All of this is meant to explicate what Sellars means by saying that "the descriptive and explanatory resources of language advance hand in hand." In addition to Kant's idea, Sellars here takes over Carnap's idea of understanding concepts whose paradigm is modal concepts as (in some sense) *metalinguistic*. The principal class of genuinely intelligible, nondefective nondescriptive vocabulary Carnap allows in *The Logical Syntax of Language* is syntactic metavocabulary and what he there calls "quasi-syntactic" vocabulary, which is covertly metalinguistic. For Sellars, the *rules* which modal vocabulary expresses are rules for deploying linguistic locutions. Their "rulishness" is their subjunctive robustness. Following out this line of thought, Sellars takes it that "grasp of a concept is mastery of the use of a word." He then understands the metalinguistic features in question in terms of rules of *inference*, whose paradigms are Carnap's L-rules and P-rules. His generic term for the inferences that articulate

the contents of ordinary empirical descriptive concepts is “material inferences.” The term is chosen to contrast with inferences that are ‘formal’ in the sense of depending on *logical* form. In another early essay he lays out the options he considers like this:

...we have been led to distinguish the following six conceptions of the status of material rules of inference:

- (1) Material rules are as essential to meaning (and hence to language and thought) as formal rules, contributing to the architectural detail of its structure within the flying buttresses of logical form.
- (2) While not essential to meaning, material rules of inference have an original authority not derived from formal rules, and play an indispensable role in our thinking on matters of fact.
- (3) Same as (2) save that the acknowledgment of material rules of inference is held to be a dispensable feature of thought, at best a matter of convenience.
- (4) Material rules of inference have a purely derivative authority, though they are genuinely rules of inference.
- (5) The sentences which raise these puzzles about material rules of inference are merely abridged formulations of logically valid inferences. (Clearly the distinction between an inference and the formulation of an inference would have to be explored).
- (6) Trains of thought which are said to be governed by "material rules of inference" are actually not inferences at all, but rather activated associations which mimic inference, concealing their intellectual nudity with stolen "therefores".<sup>47</sup>

His own position is that an expression has conceptual content conferred on it by being caught up in, playing a certain role in, material inferences:

---

<sup>47</sup> Sellars, "Inference and Meaning" *PPPW* pp. 265/317, reprinted in *In the Space of Reasons*.

...it is the first (or "rationalistic") alternative to which we are committed. According to it, material transformation rules determine the descriptive meaning of the expressions of a language within the framework provided by its logical transformation rules... In traditional language, the "content" of concepts as well as their logical "form" is determined by the rules of the Understanding.<sup>48</sup>

By "traditional language" here, he means Kantian language. The talk of "transformation rules" is, of course, Carnapian. In fact in this essay Sellars identifies his "material rules of inference" with Carnap's "P-rules." 'Determine' is crucially ambiguous between 'constrain' and 'settle'—the difference corresponding to that between what I have elsewhere called 'weak' and 'strong' semantic inferentialism.

As already indicated, the material inferential rules that in one or another of these senses "determine the descriptive meaning of expressions" are for Sellars just the subjunctively robust, hence explanation-supporting ones. As he puts the point in the title of a long essay, he construes "Concepts as Involving Laws, and Inconceivable without Them." This is his response to Quine's implicit challenge in "Two Dogmas of Empiricism" to say what feature of their use distinguishes inferences determining conceptual contents from those that simply register matters of fact. Since empirical inquiry is generally required to determine what laws govern concepts such as copper, temperature, and mass, Sellars accepts the consequence that it plays the role not only of determining facts but also of improving our conceptions—of teaching us more about the concepts that articulate those facts by teaching us more about what really follows from what.

On this way of understanding conceptual content, the modal concepts that express the lawfulness of connections among concepts and so underwrite subjunctively robust implications—concepts such as law, necessity, and what is expressed by the use of the subjunctive mood—have a different status from those of ordinary empirical descriptive concepts. Rather than in the first instance describing how the world is, they make explicit features of the framework that makes such description possible. Because they play this distinctive framework-

---

<sup>48</sup> Sellars, "Inference and Meaning" *PPPW* pp. 284/336.

explicating role, what they express must be implicitly understood by anyone who can deploy *any* ground-level descriptive concepts. As I would like to put the point, in knowing how to (being able to) use any ordinary empirical descriptive vocabulary, each interlocutor already knows how to do everything she needs to know how to do in order to be able to deploy the modal locutions that register the subjunctive robustness of the inferences that determine the content of the descriptive concepts that vocabulary expresses. This is what Kant's idea that the pure concepts of the understanding are knowable *a priori* becomes when transposed into Sellars's framework.

The two lines of thought that orient Sellars's treatment of alethic modality, semantic inferentialism and a metalinguistic understanding of the expressive role characteristic of modal locutions, are epitomized in an early formulation:

I shall be interpreting our judgments to the effect that A causally necessitates B as the expression of a rule governing our use of the terms 'A' and 'B',<sup>49</sup>

where the rule in question is understood as a rule licensing subjunctively robust inferences. I have been filling in the claim that this overall approach to modality deserves to count as a development of Kant's notion of categories, pure concepts of the understanding, as concepts that make explicit features of the discursive framework that makes empirical description possible. Sellars himself, however, does not discuss this aspect of his work under that heading. When he talks about categories he turns instead to his nominalism about abstract entities. The central text here is "Towards a Theory of the Categories" of 1970.<sup>50</sup> The story he tells there begins with Aristotle's notion of categories (though he waves his hands wistfully at a discussion of its origins in Plato's *Sophist* that he feels cannot shoehorn into the paper) as ontological *summa genera*. There he opposes an unobjectionable hierarchy

Fido is a dachshund.

Fido is a dog.

---

<sup>49</sup> Sellars, "Language, Rules, and Behavior" footnote 2 to p. 136/296 in *PPPW*.

<sup>50</sup> In *Experience and Theory*, edited by L. Foster and J.W. Swanson (University of Massachusetts Press, 1970), pp. 55-78. Reprinted in *Essays in Philosophy and its History* (D. Reidel, 1974).



Fido is a brute.

Fido is an animal.

Fido is a corporeal substance.

Fido is a substance.

To a potentially problematic one

X is a red.

X is a color.

X is a perceptual quality.

X is a quality.<sup>51</sup>

The next decisive move in understanding the latter hierarchy he attributes to Ockham, whom he reads as transposing the discussion into a metalinguistic key. Ockham's strategy, he tells us, is to understand

(A) Man is a species.

as

(B) ·Man· is a sortal mental term.<sup>52</sup>

while construing mental items as “analogous to linguistic expressions in overt speech.”

This sketch sets up the transition to what Sellars makes of Kant's understanding of categories:

What all this amounts to is that to apply Ockham's strategy to the theory of categories is to construe categories as classifications of conceptual items. This

---

<sup>51</sup> “Towards a Theory of the Categories” (TTC) §10-11.

<sup>52</sup> TTC §16.

becomes, in Kant’s hands, the idea that categories are the most generic functional classifications of the elements of judgments.<sup>53</sup>

At the end of this development from Aristotle through Ockham to Kant, he concludes

[I]nstead of being *summa genera* of entities which are objects ‘in the world,’ ...categories are *summa genera* of conceptual items.<sup>54</sup>

The account he goes on to expound in this essay, as well as in his other expositions of his nominalism about terms for qualities or properties, construes such terms metalinguistically, as referring to the inferential roles of the base-level concepts as used in empirical descriptions. I explain how I understand the view and the arguments on this topic in Chapter Seven: “Sellars’s Metalinguistic Expressive Nominalism.” Without going into that intricate view further here, the point I want to make is that although Sellars does not say so, the metaconceptual role he here explicitly puts forward as a successor-concept to Kant’s notion of category is generically the same as that I have argued he takes alethic modal locutions to play. It is this capacious conception I want to build upon and develop further.

#### 4. Categories Today

The general conception of pure categorial concepts that I have been attributing to Sellars, based on the commonalities visible in his treatment of alethic modal vocabulary and of abstract ontological vocabulary, develops Kant’s idea by treating some vocabularies (and the concepts they express) as “covertly metalinguistic.” This Sellarsian conception represents his development of Carnap’s classification of some expressions as “quasi-syntactic.” The underlying insight is that some important kinds of vocabularies that are not strictly or evidently metalinguistic are used not (only) to describe things, but in ways that (also) depend on the use of *other* vocabularies—paradigmatically, empirical descriptive ones.

---

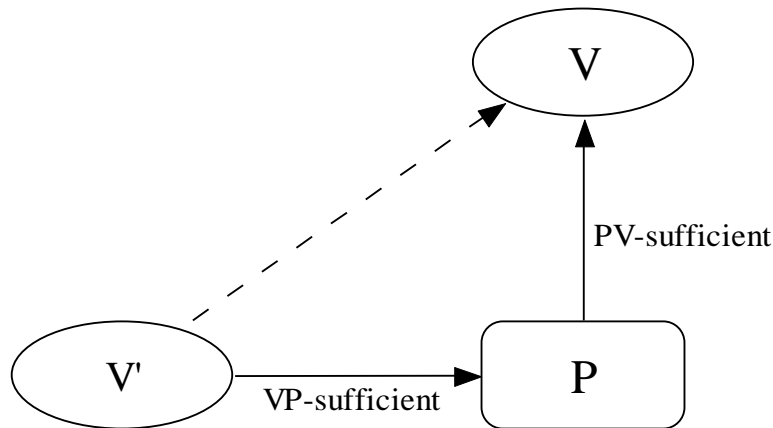
<sup>53</sup> TTC §22.

<sup>54</sup> TTC §23.

The lessons I draw from the strengths and weaknesses of Sellars's successor-conception of the "pure concepts of the Understanding" are four-fold. That is, I think he is pointing towards an expressive role characteristic of some concepts and the vocabularies expressing them that has four distinctive features. First, these concepts express what I will call "pragmatically mediated semantic relations" between vocabularies. Second, these concepts play the expressive role of making explicit essential features of the use of some other vocabulary. Third, the proper use of these concepts can be systematically elaborated from the use of that other vocabulary. Fourth, the features of vocabulary(concept)-use they explicate are universal: they are features of any and every autonomous discursive practice. I think there are concepts that play this distinctive four-fold expressive role, and that a good thing to mean today by the term "category" is metaconcepts that do so.

Carnap and Tarski introduced the expression "metalanguage" for languages that let one talk about languages, with the examples of syntactic and semantic metalanguages. In his earliest writings, Sellars also talks about "pragmatic metalanguages," meaning languages for talking about the *use* of languages—rather than the syntactic or semantic properties of expressions. These were to be the languages in which we conduct what he called "pure pragmatics." During and after Sellars's most important work in the the *anni mirabiles* of 1954-63, however (possibly influenced by Carnap), he shifts to using the expression "semantics" to cover the essentially the same ground. I think that this was a step backward, and that it is one of the obstacles that prevented him from getting clear about the sense in which he wanted to claim that such locutions as alethic modal vocabulary and singular terms purporting to refer to universals ("circularity") and their kinds ("property") are "covertly metalinguistic." One vocabulary serving as a pragmatic metavocabulary for another is the most basic kind of pragmatically mediated semantic relation between vocabularies. It deserves to be called such because the *semantics* of the pragmatic metavocabulary depends on the *use* of the vocabulary for which it is a pragmatic metavocabulary. The relation itself is aptly called a "semantic" relation in the special case where one vocabulary is sufficient to specify practices or abilities whose exercise is sufficient to confer on another vocabulary the meanings that it expresses.

We could represent such a semantic relation, mediated by the practices of using the second vocabulary that the first vocabulary specifies, like this:<sup>55</sup>

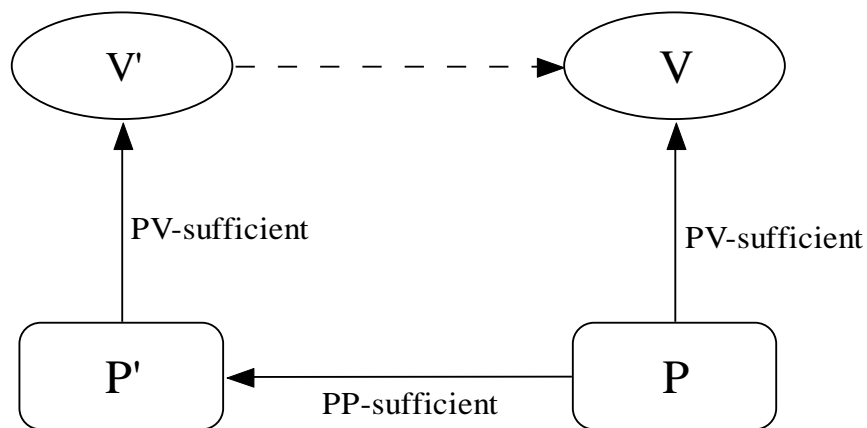


The pragmatically mediated semantic relation between vocabularies  $V'$  and  $V$ , indicated by the dashed arrow, obtains when vocabulary  $V'$  is expressively sufficient to *specify* practices-or-abilities  $P$  (that semantic fact about  $V'$  with respect to  $P$  is here called “VP-sufficiency”) that are sufficient to *deploy* the vocabulary  $V$  with the meanings that it expresses when so used. In asserting that this relation between vocabularies obtains, one is claiming that if all the sentences in  $V'$  used to specify the practices-or-abilities  $P$  are true of  $P$ , then anyone engaging in those practices or exercising those abilities as specified in  $V'$  is using the expressions of  $V$  with their proper meanings. This semantic relation between what is expressible in the two vocabularies is mediated by the practices  $P$  that the first specifies and which are the use of the second. This particular pragmatically mediated semantic relation holds when the vocabulary  $V'$  allows one to *say* what one must *do* in order to *say* what can be said in the vocabulary  $V$ . In that sense  $V'$  makes *explicit* (sayable, claimable) the practices-or-abilities *implicit* in using  $V$ . This is the explicative relation I mention as the second component of the complex expressive role that I am offering as a candidate for a contemporary successor-(meta)concept to Kant’s (meta)concept of category. There are other pragmatically mediated semantic relations besides being a pragmatic metavocabulary in this sense, and others are involved in the categorial expressive role. The

<sup>55</sup> I introduce, develop, and apply these “meaning-use diagrams” in *Between Saying and Doing: Towards an Analytic Pragmatism* [Oxford University Press, 2008].

result will still fall under the general rubric that is the first condition: being a pragmatically mediated semantic relation.

One such further pragmatically mediated semantic relations between vocabularies holds when the practices PV-sufficient for deploying one vocabulary, though not themselves PV-sufficient for deploying a second one, can be systematically elaborated into such practices. That is, in being able to deploy the first vocabulary, one already knows how to do everything one needs to know how to do, in principle, to deploy the second. But those abilities must be suitably recruited and recombined. The paradigm here is *algorithmic* elaboration of one set of abilities into another. Thus, in the sense I am after, the capacities to do multiplication and subtraction are algorithmically elaborable into the capacity to do long division. *All* you need to learn how to do is to put together what you already know how to do in the right way—a way that can be specified by an algorithm. The diagram for this sort of pragmatically mediated semantic relation between vocabularies is:



The dotted arrow indicates the semantic relation between vocabularies V' and V. It is the relation that holds when all the relations indicated by solid arrows hold—that is, when the practices-or-abilities sufficient to deploy vocabulary V can be elaborated into practices sufficient to deploy vocabulary V'. In this case, the semantic relation in question is mediated by two sets of practices-or-abilities: those sufficient to deploy the two vocabularies.

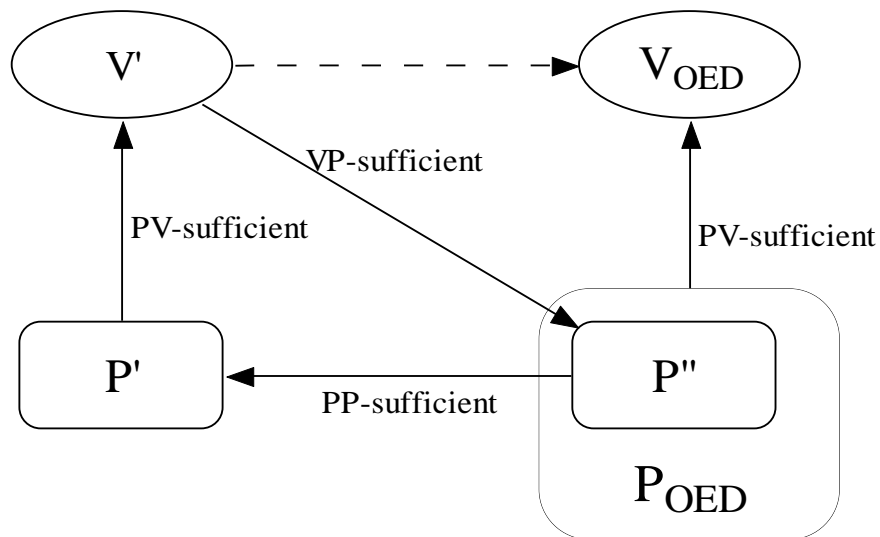
A concrete example of vocabularies standing in this pragmatically mediated semantic relation, I claim, is that of *conditionals* in relation to ordinary empirical descriptive (OED) vocabulary. For using such OED vocabulary, I claim (following Sellars following Kant), requires distinguishing in practice between materially good inferences involving descriptive predicates and ones that are not materially good. One need not be either infallible or omniscient in this regard, but unless one makes *some* such distinction, one cannot count as deploying the OED vocabulary in question. But in being able practically to distinguish (however fallibly and incompletely) between materially good and materially bad inferences, one knows how to do everything one needs to know how to do, in principle, to deploy conditionals. For conditionals can be introduced by recruiting those abilities in connection with the use of sentences formed from the old vocabulary by using the new vocabulary. On the side of circumstances of application (assertibility conditions), one must acknowledge commitment to the conditional  $p \rightarrow q$  just in case one takes the inference from  $p$  to  $q$  to be a materially good one. And on the side of consequences of application, if one acknowledges commitment to the conditional  $p \rightarrow q$ , then one must take the inference from  $p$  to  $q$  to be a materially good one. These rules constitute an algorithm for elaborating the ability to distinguish materially good from materially bad inference using OED vocabulary (or any other vocabulary, for that matter) into the ability appropriately to use conditionals formed from that vocabulary: to distinguish when such conditionals are assertible, and what the consequences of their assertibility is.

My idea for a successor-concept to what Sellars (with hints from Carnap) made of Kant's metaconception of pure concepts of the Understanding is that they must play *both* of these expressive roles, stand in *both* sorts of pragmatically mediated semantic relations to another vocabulary. It must be possible to *elaborate* their use from the use of the index vocabulary, and they must *explicate* the use of that index vocabulary. Speaking more loosely, we can say that such concepts are both *elaborated from* and *explicative of* the use of other concepts—in short that they are el-ex, or just LX with respect to the index vocabulary.

The fourth condition I imposed above is that the concepts in question must be *universally* LX, by which I mean that they must be LX for every autonomous discursive practice (ADP)—

every language game one could play though one played no other. That is, the practices from which their use can be elaborated and of which their use is explicative must be essential to talking or thinking at all. This universality would distinguish categorial concepts, in the sense being specified, from metaconcepts that were elaborated from and explicative of only some parasitic fragment of discourse—culinary, nautical, or theological vocabulary, for instance. I take it that any autonomous discursive practice must include the use of ordinary empirical descriptive vocabulary. If so, being LX for OED vocabulary would suffice for being *universally* LX, LX for every ADP.

Putting all these conditions together yields the following diagram of the pragmatically mediated semantic relation between vocabularies that obtains when vocabulary V' plays the expressive role of being universally LX by being elaboratable from and explicative of practices necessary for the deployment of ordinary empirical descriptive vocabulary:



The fact that the rounded rectangle labeled P'', representing the practices from which vocabulary V' is elaborated and of which it is explicative, appears inside the rounded rectangle representing practices sufficient to deploy ordinary empirical descriptive vocabulary indicates that the practices P'' are a necessary part of the practices sufficient to deploy OED vocabulary, but need not comprise all such practices. Thus, distinguishing materially good from materially bad inferences involving them is necessary for deploying ordinary empirical descriptive vocabulary

(rather than mere labels), but there is a lot more involved in doing so—using such vocabulary observationally, for instance. Different categorial metaconcepts can be LX for different essential features of the use of empirical descriptive vocabulary. Thus alethic modal vocabulary explicates the subjunctive robustness of the inferences explicated by conditionals. “Quasi-syntactic” abstract ontological vocabulary such as ‘property’ and ‘proposition’ explicate structural features of descriptive sentences.

Diagramming the expressive role of being LX for practices necessary to deploy OED vocabulary provides an analysis that breaks down the claim that some vocabulary plays a categorial role into its component sub-claims. To show that alethic modal vocabulary, for instance, stands in this pragmatically mediated semantic relation to ordinary empirical descriptive vocabulary one must show that there is some practices-or-abilities (in this case, to reason subjunctively or counterfactually) that are 1) a necessary component of practices-or-abilities that are 2) (PV)sufficient to deploy OED vocabulary, 3) from which one can elaborate practices-or-abilities that are 4) (PV)sufficient to deploy vocabulary (alethic modal vocabulary) 5) that is (VP)sufficient to explicate or specify the original practices-or-abilities. Although there is by design considerable elasticity in the concepts vocabulary, practices-or-abilities, and the various sufficiency and necessity relations between them, the fine structure of the distinctive expressive role in question is clearly specified.

What credentials does that expressive role have to pick out a worthy successor metaconcept to what Sellars made of Kant’s categories or pure concepts of the Understanding? At the beginning of my story I introduced the idea behind the Kantian categories as the idea that besides the concepts whose principal use is in giving empirical descriptions and explanations, there are concepts whose principal use is in making explicit features of the framework that makes empirical description and explanation possible. The expressive task characteristic of concepts of this latter class is to articulate what Kant called the “transcendental conditions of experience.” The concepts expressed by vocabularies that are LX for empirical descriptive vocabulary perform this defining task of concepts that are categories. As explicative of practices necessary for deploying vocabularies performing the complex expressive task of description and



explanation (distinguishable only in the context of their complementary relations within a pragmatic and semantic context that necessarily involves both), this kind of vocabulary makes it possible to *say* what practitioners must be able to *do* in order to describe and explain how things empirically are. They do this by providing a pragmatic metavocabulary for describing and explaining. This is a central feature (the ‘X’ in ‘LX’) of the complex pragmatically mediated semantic relation between categorial metaconcepts and ordinary empirical descriptive vocabulary.

One feature of the concepts performing this explicative function that Kant emphasizes is that they are “*pure* concepts of the Understanding.” (I take it that the “of” should be understood as expressing both the subjective and objective genitives—as in “Critique of Pure Reason.” These concepts both belong to the Understanding and address it, being both discursive and metaconceptual.) To say that they are pure concepts is to say that they are graspable *a priori*.<sup>56</sup> The feature of the LX model that corresponds to the a prioricity of Kant’s categories is that the use of LX metaconcepts can be elaborated from that of the empirical descriptive vocabularies for which they are LX. As I have put the point, in knowing how to deploy OED vocabulary, one already knows how to do everything one needs to know how to do to deploy vocabulary that is LX for it—such as alethic modal vocabulary, conditionals, and ontological classificatory vocabulary. If we take it, as per Sellars, that grasp of a concept is mastery of the use of a word, then one need not *actually* grasp concepts that are LX for descriptive vocabulary in order to deploy descriptive vocabulary. But in effect, *all* one is missing is the words for them. The circumstances and consequences of application of LX concepts can be formulated by rules that appeal only to abilities one already has in virtue of being able to use OED vocabulary. (Think of the sample rules for conditionals sketched above.) In that sense, the LX concepts are *implicit in* the descriptive concepts. It is not that one must or could grasp these concepts *before* deploying descriptive concepts. It is rather that nothing more is required to grasp them than is required to deploy descriptive concepts, and there are no particular descriptive concepts one must be able to

---

<sup>56</sup> Kant does admit also impure *a priori* principles.

deploy, nor any particular descriptive claims that one must endorse, in order to possess abilities sufficient to deploy the universally LX metaconcepts.

The class of concepts that are arguably universally LX (LX for every autonomous discursive practice because LX for OED vocabulary) overlaps Kant's categories in important ways—most notably in the alethic modal concepts that make explicit subjunctively robust consequential relations among descriptive concepts. But the two do not simply coincide. In *Between Saying and Doing* I argue that besides modal vocabulary, logical vocabulary, indexical and demonstrative vocabulary, normative vocabulary, and semantic and intentional vocabulary all should be thought of as LX for OED vocabulary. In spite of this extensional divergence, the fact that vocabulary that is LX for descriptive vocabulary in general principle shares with Kant's categories the two crucial features of being explicative of such vocabulary and being graspable *a priori* makes the idea of universally LX metaconcepts a worthy successor to Kant's breakthrough idea. The fact that Sellars's own development of this idea of Kant's takes such important steps in this direction convinces me that his version of the categories was a progressive step, and a Good Idea.

## 2014 Nordic Pragmatism Lectures:

### Analytic Pragmatism, Expressivism, and Modality

#### Lecture 3

#### Towards an Analytic Pragmatism: Meaning-Use Analysis

Within the Anglophone tradition, pragmatism has often appeared as a current of thought that stands apart from, and indeed runs in opposition to, the mainstream of analytic philosophy. This is true whether one uses ‘pragmatist’ in a narrow sense tailored to the triumvirate of Peirce, James, and Dewey (here one might think of Russell’s dismissive responses to the latter two), or in a more capacious sense that includes the early Heidegger, the later Wittgenstein, and, more recently, neo-pragmatists such as Rorty and Putnam. There are good reasons on both sides for adopting somewhat adversarial stances, but I think that when we examine them more closely it becomes possible to see the outlines of a common project, in the service of which the two camps might find themselves joining forces. In my 2006 John Locke lectures, entitled *Between Saying and Doing: Towards an Analytic Pragmatism*<sup>57</sup>, I explore in more detail one way of pursuing such a project. In this essay I want to offer a sketch of the basic understanding of the principal aims of the two movements that motivates that more extended discussion, and to indicate in general terms the sort of *pragmatic* semantic analysis (not, I will be insisting, an oxymoron) that might emerge from unifying their only apparently disparate concerns. The intended spirit is irenic, synthetic, and constructive.

---

<sup>57</sup> Oxford: Oxford University Press, 2008.

## I. The Classical Project of Analysis

I think of analytic philosophy as having at its center a concern with semantic relations between what I will call ‘vocabularies’. Its characteristic form of question is whether and in what way one can make sense of the meanings expressed by *one* kind of locution in terms of the meanings expressed by *another* kind of locution. So, for instance, two early paradigmatic projects were to show that everything expressible in the vocabulary of number-theory, and again, everything expressible using definite descriptions, is expressible already in the vocabulary of first-order quantificational logic with identity.

The nature of the key kind of semantic relation between vocabularies has been variously characterized during the history of analytic philosophy: as analysis, definition, paraphrase, translation, reduction of different sorts, truth-making, and various kinds of supervenience—to name just a few contenders. In each case, however, it is characteristic of classical analytic philosophy that *logical* vocabulary is accorded a privileged role in specifying these semantic relations. It has always been taken at least to be *licit* to appeal to logical vocabulary in elaborating the relation between *analysandum* and *analysans*—target vocabulary and base vocabulary. I will refer to this aspect of the analytic project as its commitment to ‘*semantic logicism*’.<sup>58</sup>

If we ask *which* were the vocabulary-kinds whose semantic relations it was thought to be important to investigate during this period, at least two *core programs* of classical analytic

---

58 In this usage, the logicism about mathematics characteristic of Frege’s *Grundgesetze* and Russell and Whitehead’s *Principia* is semantic logicism about the relations between mathematical and logical vocabularies.

philosophy show up: *empiricism* and *naturalism*. These venerable modern philosophical traditions in epistemology and ontology respectively were transformed in the twentieth century, first by being transposed into a *semantic* key, and second by the application of the newly available *logical* vocabulary to the self-consciously semantic programs they then became.

As *base* vocabularies, different species of *empiricism* appealed to phenomenal vocabulary, expressing how things appear, or to secondary-quality vocabulary, or, less demandingly, to observational vocabulary. Typical *target* vocabularies include objective vocabulary formulating claims about how things actually are (as opposed to how they merely appear), primary-quality vocabulary, theoretical vocabulary, and modal, normative, and semantic vocabularies. The generic challenge is to show how what is expressed by the use of such target vocabularies can be reconstructed from what is expressed by the base vocabulary, when it is elaborated by the use of logical vocabulary.

As *base* vocabularies, different species of *naturalism* appealed to the vocabulary of fundamental physics, or to the vocabulary of the natural sciences (including the special sciences) more generally, or just to objective descriptive vocabulary, even when not regimented by incorporation into explicit scientific theories. Typical targets include normative, semantic, and intentional vocabularies.

## II. The Pragmatist Challenge

What I want to call the “classical project of analysis,” then, aims to exhibit the meanings expressed by various target vocabularies as intelligible by means of the logical elaboration of the meanings expressed by base vocabularies thought to be privileged in some important respects—epistemological, ontological, or semantic—relative to those others. This enterprise is visible in its purest form in what I have called the “core programs” of empiricism and naturalism, in their various forms. In my view the most significant conceptual development in this tradition—the biggest thing that ever happened to it—is the *pragmatist challenge* to it that was mounted during the middle years of the twentieth century. Generically, this movement of thought amounts to a displacement from the center of philosophical attention of the notion of meaning in favor of that of use: in suitably broad senses of those terms, replacing concern with *semantics* by concern with *pragmatics*. The towering figure behind this conceptual sea-change is of course, Wittgenstein. In characterizing it, however, it will be useful to approach his radical and comprehensive critique by means of some more local, semantically corrosive argumentative appeals to the *practices of deploying* various vocabularies rather than the *meanings* they express.

Wilfrid Sellars (one of my particular heroes) criticizes the empiricist core program of the classical project of analysis on the basis of what one must *do* in order to *use* various vocabularies, and so to count as *saying* or *thinking* various things. He argues that *none* of the various candidates for empiricist base vocabularies are practically *autonomous*, that is, could be deployed in a language game one played though one played no other. For instance, no discursive practice can consist entirely of making non-inferential observation reports. For such reliably differentially elicited responses qualify as *conceptually* contentful or *cognitively* significant only insofar as they can serve as *premises* from which it is appropriate to draw *conclusions*, that is, as *reasons* for other judgments. Drawing such conclusions is applying concepts *inferentially*—that is, precisely *not* making non-inferential *observational* use of them.<sup>59</sup>

---

<sup>59</sup> This argument occupies roughly the first half of his classic *Empiricism and the Philosophy of Mind*, reprinted with an introduction by Richard Rorty and a study guide by Robert Brandom (Cambridge, MA: Harvard University Press, 1997). His critique of the phenomenalist version of empiricism can be found in “Phenomenalism,” in *In the*

Quine offers an even broader pragmatist objection, not only to the empiricist program, but to essential aspects of the whole analytic semantic project. For he attacks the very notion of meaning it presupposes. Quine is what in Chapter Two I called a “methodological” pragmatist. That is, he takes it that the whole point of a theory of meaning is to explain, codify, or illuminate features of the *use* of linguistic expressions. He, like Dummett, endorses the analogy: *meaning* is to *use* as *theory* is to *observation*. And he argues that postulating meanings associated with bits of vocabulary yields a *bad* theory of discursive practice.

If there were such things as meanings that determine how it would be correct to use our expressions, then those meanings would at least have to determine the inferential roles of those expressions: what follows from applying them, what applying them rules out, what is good evidence for or against doing so. But what follows from what depends on what else is true—on laws of nature and obscure contingent facts—that is, on what claims can serve as auxiliary hypotheses or collateral premises in those inferences. If we look at what practical abilities are required to deploy various bits of vocabulary—at what one has to be able to *do* in order to count as *saying* something with them—we do not find any special set of these whose practical significance can be understood as *pragmatically* distinctive of *semantically* necessary or sufficient conditions.

Quine thought one could save at least the naturalist program by retreating semantically to the level of reference and truth-conditions. James and Dewey appeal to the same sort of methodological pragmatism in support of more sweeping sorts of semantic revisionism—pursuing programs that Rorty, for instance, argues should be understood as more rejectionist than properly revisionist. And under the banner “Don’t look to the meaning, look to the use,” Wittgenstein further radicalizes the pragmatist critique of semantics. Pointing out to begin with that one cannot assume that uses of singular terms have the job of picking out objects, nor that

---

*Space of Reasons*, ed. Kevin Scharp and Robert Brandom (Cambridge, MA: Harvard University Press, 2007), p. 303-49.

declarative sentences are in the business of stating facts, he goes on to deny, in effect, that such uses even form a privileged center, on the basis of which one can understand more peripheral ones. (“Language,” he says, “has no downtown.”)

I take it that Wittgenstein also takes the home language game of the concept of meaning to be explanation of how expressions are correctly *used*. And he is profoundly skeptical about the utility or applicability of the model of postulation, explanation, and theoretical systematization in the case of discursive practices—about the possibility of systematically *deriving* aspects of correct use from assigned meanings. Seen from this perspective, the idea of the classical project of analysis is to codify, using logical vocabulary, the meanings expressed by one vocabulary—from which we are to derive proprieties of its use—from the meanings expressed by some *other* vocabulary—from which we can derive proprieties of *its* use. One idea, I think, is that this enterprise makes sense only if we think of the uses as species of a genus—of them all being the same general *kind* of use, say describing, stating facts, or representing states of affairs. This may seem plausible if we focus on a very restricted set of uses—just as, in the case of tools, we might be impressed to notice that nails and hammer, screws and screwdriver, glue and brush all have the function of attaching more-or-less flat things to one another. So we can think of declarative sentences as stating empirical, physical, normative, modal, and intentional facts, making *claims* about such states of affairs (even if we then find ourselves metaphysically puzzled about the nature of the fact-kinds to which we have thereby committed ourselves). But if we think of the uses as *very* different, if we think also about the carpenter’s level, pencil, and tool-belt, if we think of linguistic practice as a *motley*, of uses as not coming in a simple, or systematic, or even determinate variety, then the very idea that there is such a thing as *meanings* that permit the codification of proprieties of quite disparate kinds of use—even with liberal use of logical elaboration of the meanings—becomes contentious and in need of justification both in general and in each particular case.

More specifically, Wittgenstein uses the image of “family resemblances” to urge that the *kinds* into which linguistic practices and the vocabularies caught up in them are functionally sorted—what belong together in boxes labeled ‘game’, ‘name’, ‘description’, ‘assertion’,



‘observation’ and so on—do not typically admit of specification in terms of underlying principles specifiable in other vocabularies, whether by genus and differentia(e) or any other kind of explicit rule or definition. It is easy to understand this line of thought as entailing a straightforward denial of the possibility of semantic analysis in the classical sense.

I think that one thought underlying these observations about the unsystematic, unsurveyable variety of kinds of uses of expressions and about the uncodifiable character of those kinds concerns the essentially *dynamic* character of linguistic practice. I think Wittgenstein thinks that an absolutely fundamental discursive phenomenon is the way in which the abilities required to deploy one vocabulary can be practically *extended*, elaborated, or developed so as to constitute the ability to deploy some further vocabulary, or to deploy the old vocabulary in quite different ways. Many of his thought-experiments concern this sort of process of *pragmatic projection* of one practice into another. We are asked to imagine a community that uses proper names only for people, but then extends the practice to include rivers. There is no guarantee that interlocutors can master the extended practice, building on what they can already do. But if they can, then they will have changed the only ‘essence’ proper-name usage could be taken to have had.<sup>60</sup> In the old practice it always made sense to ask for the identity of the *mother* and *father* of the named item; in the new practice, that question is often senseless. Again, we are asked to imagine a community that talked about having gold or silver in one’s teeth, and extends that practice to talk about having pain in one’s teeth. If as a matter of contingent fact the practitioners can learn to use the expression ‘in’ in the new way, building on but adapting the old, they will have fundamentally changed the ‘meaning’ of ‘in’. In the old practice it made sense to ask where the gold was *before* it was in one’s tooth; in the new practice asking where the pain was before it was in the tooth can lead only to a distinctively *philosophical* kind of puzzlement.<sup>61</sup>

---

60 Cf. Quine’s remark (in “Two Dogmas of Empiricism”): “Meaning is what essence becomes when it is detached from the thing and attached to the word.” I distinguish ‘scare quotes’ from regular quotes by using superscripted ‘s’s. My official theory of such quotation, as the conceptual converse of *de re* ascriptions of propositional attitude, can be found in *Making It Explicit*, pp. 545-7 and 588-90.

61 I am indebted for this way of thinking of Wittgenstein’s point to Hans Julius Schneider’s penetrating discussion in his *Phantasie und Kalkul* (Frankfurt: Suhrkamp, 1992).

At every stage, what practical extensions of a given practice are possible for the practitioners can turn on features of their embodiment, lives, environment, and history that are contingent and wholly particular to them. And which of those developments actually took place, and in what order can turn on any obscure fact. The reason vocabulary-kinds resist specification by rules, principles, definitions, or meanings expressed in other vocabularies is that they are the current time-slices of processes of development of practices that have this dynamic character—and that is why the collection of uses that is the current cumulative and collective result of such developments-by-practical-projection is a motley.<sup>62</sup> If that is right, then any codification or theoretical systematization of the uses of those vocabulary-kinds by associating with them meanings that determine which uses are correct will, if at all successful, be successful only contingently, locally, and temporarily. Semantics on this view is an inherently Procrustean enterprise, which can proceed only by *theoretically* privileging some aspects of the use of a vocabulary that are not at all *practically* privileged, and spawning philosophical puzzlement about the intelligibility of the rest.<sup>63</sup> On this conception, the classical project of analysis is disease that rests on a fundamental, if perennial, misunderstanding—one that can be removed or ameliorated only by heeding the advice to replace concern with *meaning* by concern with *use*. The recommended philosophical attitude to discursive practice is accordingly *descriptive particularism, theoretical quietism, and semantic pessimism*.

---

62 A patient and detailed investigation of the mechanisms of this phenomenon in basic descriptive and scientific concepts, and an extended argument for its ubiquity can be found in Mark Wilson's exciting and original *Wandering Significance* (Oxford: Oxford University Press, 2006).

63 I would be happy if those who dance with his texts find affinities here with Hegel's insistence that the metaconceptual categories of *Verstand* must be replaced by those of *Vernunft*. It is characteristic of his philosophical ambition that Hegel draws the opposite of Wittgenstein's conclusions from an appreciation of the dynamics of conceptual development and its sensitivity to arbitrary contingent features of the practitioners, devoting himself to elaborating what he insists is the *logic* of such processes and the conceptual contents they shape.

### III. Extending the Project of Analysis: Pragmatically Mediated Semantic Relations

On this account Wittgenstein is putting in place a picture of discursive meaningfulness or significance that is very different from that on which the classical project of analysis is predicated. In place of *semantics*, we are encouraged to do *pragmatics*—not in the sense of Kaplan and Stalnaker, which is really the semantics of token-reflexive expressions, nor again in the sense of Grice, which addresses conversational heuristics in terms that presuppose a prior, independent, classical semantics—but ‘pragmatics’ in the sense of the study of the *use* of expressions in virtue of which they are meaningful at all. To the formal, mathematically inspired tradition of Frege, Russell, Carnap, and Tarski, culminating in model-theoretic and possible worlds semantics, is opposed an anthropological, natural-historical, social-practical inquiry aimed both at demystifying our discursive doings, and at deflating philosophers’ systematic and theoretical ambitions regarding them. I think that contemporary philosophers of language have tended to draw this opposition in the starkest possible terms, treating these approaches as mutually exclusive, hence as requiring that a choice be made between them, thereby marking out a substantial sociological faultline in the discipline. Those who are moved by the pragmatist picture generally accept the particularist, quietist conclusions Wittgenstein seems to have drawn from it. And those committed to some version of the project of semantic analysis have felt obliged to deny the significance of pragmatics in this sense, or at the least to dismiss it as irrelevant to properly semantic concerns. In the most extreme cases, the attitudes of anti-pragmatist philosophers of language to Wittgenstein’s picture verges on that of the Victorian lady to Darwin’s theory: One hopes that it is not true, and that if it is true, at least that it not become generally known.

But I do not think we are obliged to choose between these approaches. They should be seen as complementing rather than competing with one another. Semantics and pragmatics, concern with meaning and concern with use, ought surely to be understood as aspects of one, more comprehensive, picture of the discursive. Pragmatist considerations do not oblige us to focus on pragmatics to the exclusion of semantics; we can deepen our semantics by the addition of pragmatics. If we extract consequences from the pragmatists’ observations somewhat more

modestly and construe the analytic project somewhat more broadly, the two will be seen not only as compatible, but as mutually illuminating. If we approach the pragmatists' observations in an analytic spirit, we can understand pragmatics as providing special resources for extending and expanding the analytic semantic project, from exclusive concern with relations among meanings to encompass also relations between meaning and use. In its most ambitious form, as in the present project, such an enterprise would aspire to articulate something like a *logic* of the relations between meaning and use.

If we leave open the possibility that the use of some vocabulary may be illuminated by taking it to express some sort of meaning or content—that is, if we do not from the beginning embrace theoretical semantic nihilism—then the most important positive pragmatist insight will be one complementary to the methodological pragmatism I have already identified. The thought underlying the pragmatist line of thought is that what makes some bit of vocabulary mean what it does is how it is used. What we could call *semantic* pragmatism is the view that the only explanation there could be for how a given *meaning* gets associated with a vocabulary is to be found in the *use* of that vocabulary: the practices by which that meaning is conferred or the abilities whose exercise constitutes deploying a vocabulary with that meaning. To broaden the classical project of analysis in the light of the pragmatists' insistence on the centrality of pragmatics, we can focus on this fundamental relation between use and meaning, between practices or practical abilities and vocabularies. We must look at what it is to use locutions *as* expressing meanings—that is, at what one must *do* in order to count as *saying* what the vocabulary lets practitioners express. I am going to call this kind of relation “practice-vocabulary sufficiency”—or usually, “PV-sufficiency,” for short. It obtains when engaging in a specified set of practices or exercising a specified set of abilities<sup>64</sup> is sufficient for someone to count as *deploying* a specified vocabulary.

---

64 For the purposes of the present project, I will maintain a studied neutrality between these options. The apparatus I am introducing can be noncommittal as to whether we understand content-conferring *uses* of expressions in terms of social practices or individual abilities.

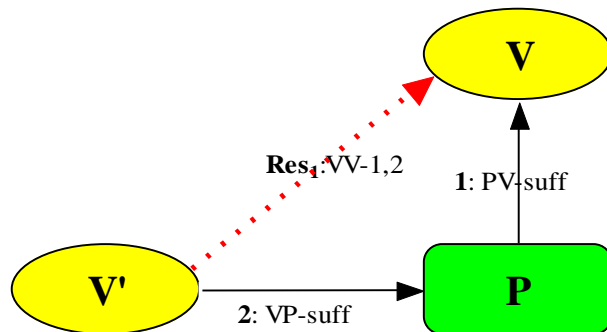
Of course it matters a lot how we think about these content-conferring, vocabulary-deploying practices or abilities. The semantic pragmatist's claim that use confers meaning (so talk of practices or the exercise of abilities as deploying vocabularies) reverts to triviality if we are allowed to talk about "using the tilde to express negation," "the ability to mean red by the word 'red'," or "the capacity to refer to electrons by the word 'electron'," (or, I think, even *intentions* so to refer). And that is to say that the interest of the PV-sufficiency of some set of practices or abilities for the deploying of a vocabulary is quite sensitive to the *vocabulary* in which we *specify* those practices-or-abilities. Talk of practices-or-abilities has a definite sense only insofar as it is relativized to the vocabulary in which those practices-or-abilities are specified. And that means that besides PV-sufficiency, we should consider a second basic meaning-use relation: "vocabulary-practice sufficiency," or just "VP-sufficiency," is the relation that holds between a vocabulary and a set of practices-or-abilities when that vocabulary is sufficient to *specify* those practices-or-abilities.<sup>65</sup> VP-sufficient vocabularies that *specify* PV-sufficient practices let one *say* what it is one must *do* to count as engaging in those practices or exercising those abilities, and so to deploy a vocabulary to *say* something.

PV-sufficiency and VP-sufficiency are two basic *meaning-use* relations (MURs). In terms of those basic relations, we can define a more complex relation: the relation that holds between vocabulary V' and vocabulary V when V' is VP-sufficient to specify practices-or-abilities P that are PV-sufficient to deploy vocabulary V. This VV-relation is the *composition* of the two basic MURs. When it obtains I will say that V' is a *pragmatic metavocabulary* for V. It allows one to *say* what one must *do* in order to count as *saying* the things expressed by vocabulary V. We can present this relation graphically in a *meaning-use diagram* (MUD):

---

65 Somewhat more precisely: some theory (a set of sentences), formulable in the vocabulary in question, is such that if all those sentences are true of some interlocutor, then it thereby counts as exercising the relevant ability, or engaging in the relevant practices.

**Meaning-Use Diagram #1:  
Pragmatic  
Metavocabulary**



The conventions of this diagram are:

- Vocabularies are shown as ovals, practices-or-abilities as (rounded) rectangles.
- Basic meaning-use relations are indicated by solid arrows, numbered and labeled as to kind of relation.
- Resultant meaning-use relations are indicated by dotted arrows, numbered, and labeled as to kind and the basic MURs from which they result.

The idea is that a resultant MUR is the relation that obtains when all of the basic MURs listed on its label obtain.

Being a pragmatic metavocabulary is the simplest species of the genus I want to introduce here. It is a *pragmatically mediated semantic relation* between vocabularies. It is pragmatically mediated by the practices-or-abilities that are *specified* by one of the vocabularies (which *say* what counts as *doing* that) and that *deploy* or are the *use* of the other vocabulary (what one says *by* doing that). The semantic relation that is established thereby between the two vocabularies is of a distinctive sort, quite different from, for instance, definability, translatability, reducibility, and supervenience. **My basic suggestion for extending the classical project of analysis so as to incorporate as essential positive elements the insights that animate the pragmatist critique of that project is that, alongside the classical semantic relations between**

**vocabularies that project has traditionally appealed to, we consider also *pragmatically mediated ones*—of which the relation of being a pragmatic metavocabulary is a paradigm.**

Under what circumstances would this simplest pragmatically mediated semantic relation be philosophically interesting, when considered in connection with the sorts of vocabularies that have traditionally been of most interest to classical analysis? At least one sort of result that could be of considerable potential significance, I think, is if it turned out that in some cases pragmatic metavocabularies exist that differ significantly in their expressive power from the vocabularies for the deployment of which they specify sufficient practices-or-abilities. I will call that phenomenon “*pragmatic expressive bootstrapping*.” If one vocabulary is strictly weaker in expressive power than the other, I will call that *strict expressive bootstrapping*. We are familiar with this sort of phenomenon in ordinary semantics, where sometimes a semantic metalanguage differs substantially in expressive power from its object language—for instance, where we can produce an extensional metalanguage for intensional languages, as in the case of possible worlds semantics for modality. One example of a claim of this shape in the case of pragmatically mediated semantic relations—though of course it is not expressed in terms of the machinery I have been introducing—is Huw Price’s pragmatic normative naturalism.<sup>66</sup> He argues, in effect, that although normative vocabulary is not *reducible to* naturalistic vocabulary, it might still be possible to *say* in wholly naturalistic vocabulary what one must *do* in order to be *using* normative vocabulary. If such a claim about the existence of an expressively bootstrapping naturalistic pragmatic metavocabulary for normative vocabulary could be made out, it would evidently be an important chapter in the development of the naturalist core program of the classical project of philosophical analysis. It would be a paradigm of the sort of payoff we could expect from extending that analytic project by including pragmatically mediated semantic relations. (Later on I’ll discuss briefly a claim of this shape concerning *indexical* vocabulary.)

---

66 See his “Naturalism without Representationalism,” in *Naturalism in Question*, ed. Mario de Caro and David Macarthur, (Cambridge, MA: Harvard University Press, 2004), pp. 71-90. Price calls his view “subject naturalism,” as opposed to the more traditional (and more metaphysical) “object naturalism.”

The meaning-use diagram of the pragmatically mediated semantic relation of being a pragmatic metavocabulary illustrates a distinctive kind of *analysis* of that relation. It exhibits that relation as the resultant, by composition, of the two basic meaning-use relations of PV-sufficiency and VP-sufficiency. A complex MUR is analyzed as the product of operations applied to basic MURs. This is *meaning-use analysis*. The same analytic apparatus applies also to more complex pragmatically mediated semantic relations. Consider one of the pragmatist criticisms that Sellars addresses to the empiricist core program of the classical analytic project. It turns on the assertion of the *pragmatic dependence* of one set of vocabulary-deploying practices-or-abilities on another.

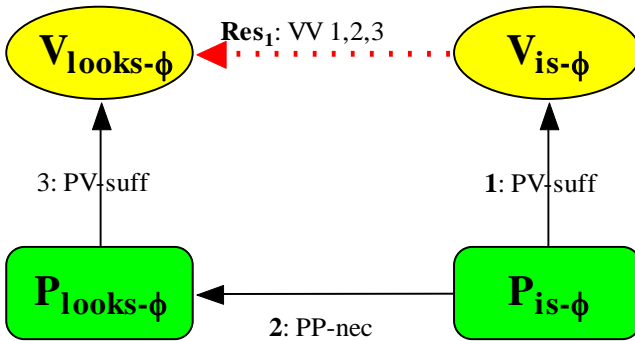
Because he thinks part of what one is *doing* in saying how things merely appear is withholding a commitment to their actually being that way, and because one cannot be understood as *withholding* a commitment that one cannot *undertake*, Sellars concludes that one cannot have the ability to say or think how things *seem* or *appear* unless one also has the ability to make claims about how things *actually are*. In effect, this Sellarsian pragmatist critique of the phenomenalist form of empiricism consists in the claim that the practices that are PV-sufficient for ‘is’- $\phi$  talk are PP-necessary for the practices that are PV-sufficient for ‘looks’- $\phi$  talk.<sup>67</sup> That pragmatic dependence of practices-or-abilities then induces a resultant pragmatically mediated semantic relation between the vocabularies. The meaning-use diagram for this claim is:

---

<sup>67</sup> I discuss this argument in greater detail in the final chapter of *Tales of the Mighty Dead* (Cambridge, MA: Harvard University Press 2004).



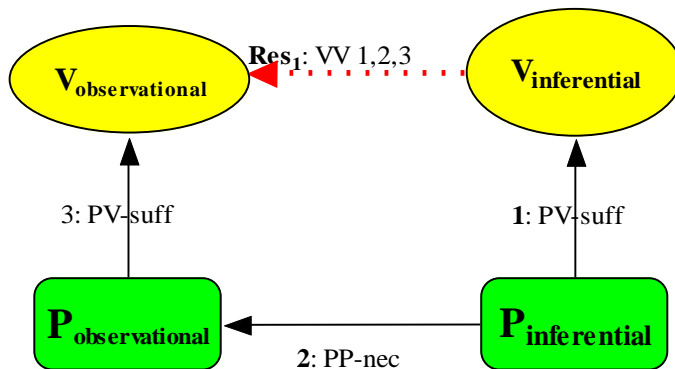
**Meaning-Use Diagram #2:  
Pragmatically Mediated  
Semantic Presupposition**



The resultant MUR here is a kind of complex, pragmatically mediated, VV-necessity, or semantic presupposition.

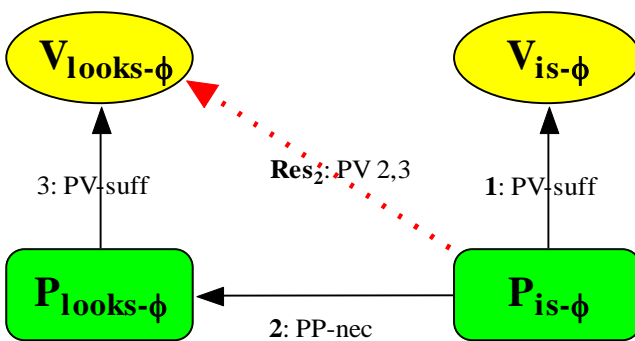
In fact, although Sellars’s argument for the crucial PP-necessity relation of pragmatic dependence of one set of vocabulary-deploying practices-or-abilities on another is different, his argument against the observational version of empiricism—the claim that purely non-inferential, observational uses do not form an autonomous discursive practice, but presuppose inferential uses—has exactly the same form:

**Meaning-Use Diagram #3:  
Pragmatically Mediated  
Semantic Presupposition**



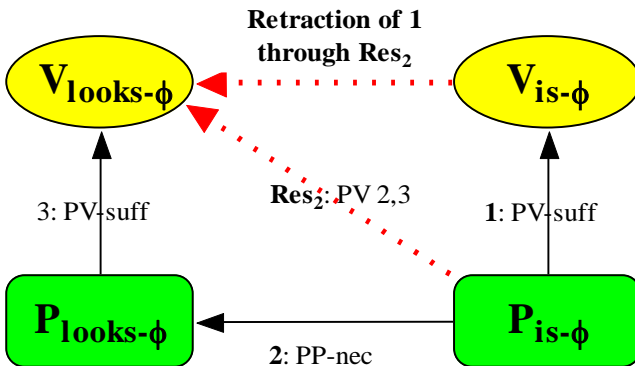
For these cases, we can say something further about the nature of the pragmatically mediated semantic relation that is analyzed as the resultant MUR in these diagrams. For instead of jumping directly to this VV resultant MUR, we could have put in the composition of the PP-necessity and second PV-sufficiency relation, yielding a kind of complex pragmatic presupposition:

**Meaning-Use Diagram #4:  
Composition**



If this diagram were completed by an arrow from  $V_{is-\phi}$  to  $V_{looks-\phi}$  such that the *same* diagonal resultant arrow could represent *both* the composition of relations 2 and 3 *and* the composition of relation 1 and the newly supplied one, then category theorists would say that the diagram *commutes*. And the arrow that needs to be supplied to make the diagram commute they call the *retraction* of relation 1 through the composition  $Res_2$ :

**Meaning-Use Diagram #5:  
Composition and  
Retraction**



After composition, then, the next most complex form of resultant MUR is retraction. Analyzing the structure of Sellars’s pragmatist arguments against empiricism requires recognizing the pragmatically mediated semantic relation he claims holds between phenomenal and objective vocabulary as the retraction of a constellation of more basic meaning-use relations.

#### IV. Automata: Syntactic PV-sufficiency and VP-sufficiency

Now this is all extremely abstract. To make it more definite, we need to fill in (at least) the notions of vocabulary, practice-or-ability, PV-sufficiency, and VP-sufficiency, which are the fundamental elements that articulate what I am calling the “meaning-use analysis” of resultant meaning-use relations—in particular, the pragmatically mediated semantic relations between vocabularies that I am claiming we must acknowledge in order to pursue the classical project of philosophical analysis in the light of what is right about the pragmatist critique of it. We can begin to do that by looking at a special case in which it is possible to be unusually clear and precise about the things and relations that play these metatheoretic roles. This is the case where ‘vocabulary’ takes a purely *syntactic* sense. Of course, the cases we eventually care about involve vocabularies understood in a sense that includes their *semantic* significance. But besides

the advantages of clarity and simplicity, we will find that some important lessons carry over from the syntactic to the semantic case.

The restriction to vocabularies understood in a spare syntactic sense leads to correspondingly restricted notions of what it is to *deploy* such a vocabulary, and what it is to *specify* practices-or-abilities sufficient to deploy one. Suppose we are given an *alphabet*, which is a finite set of primitive sign types—for instance, the letters of the English alphabet. The *universe* generated by that alphabet then consists of all the finite strings that can be formed by concatenating elements drawn from the alphabet. A *vocabulary* over such an alphabet—in the syntactic sense I am now after—is then any subset of the universe of strings that alphabet generates. If the generating alphabet is the English alphabet, then the vocabulary might consist of all English sentences, all possible English texts, or all and only the sentences of *Making It Explicit*.<sup>68</sup>

What can we say about the abilities that count as *deploying* a vocabulary in this spare syntactic sense?<sup>69</sup> The abilities in question are the capacity to *read* and *write* the vocabulary. In this purely syntactic sense, ‘reading’ it means being able practically to *distinguish* within the universe generated by the alphabet, strings that do, from those that do not, belong to the specified vocabulary. And ‘writing’ it means practically being able to *produce* all and only the strings in the alphabetic universe that do belong to the vocabulary.

We assume as primitive abilities the capacities to read and write, in this sense, the alphabet from whose universe the vocabulary is drawn—that is, the capacity to respond differentially to alphabetic tokens according to their type, and to produce tokens of antecedently

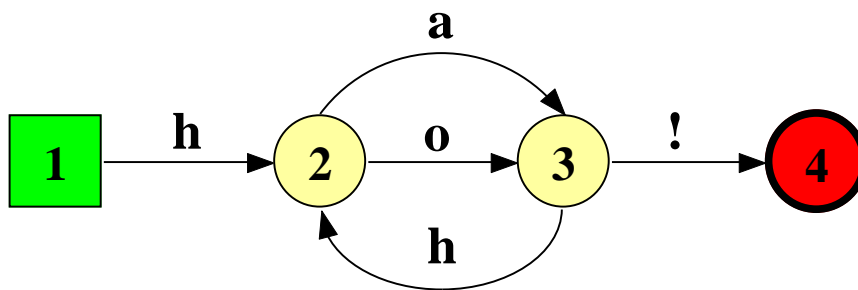
---

68 Computational linguists, who worry about vocabularies in this sense, have developed metalanguages for specifying important classes of such vocabularies: the syntactic analogues of semantic metalanguages in the cases we will eventually address. So, for instance, for the alphabet {a,b}, ‘a<sup>n</sup>b<sup>n</sup>’ characterizes the vocabulary that comprises all strings of some finite number of ‘a’s followed by the *same* number of ‘b’s. ‘a(ba)<sup>\*</sup>b’ characterizes the vocabulary that comprises all strings beginning with an ‘a’, ending with a ‘b’, and having *any* number of repetitions of the sub-string ‘ba’ in between.

69 Here we can safely just talk about *abilities*, without danger of restricting the generality of the analysis.

specified alphabetic types. Then the abilities that are PV-sufficient to deploy some vocabularies can be specified in a particularly simple form. They are *finite-state automata* (FSAs). As an example, suppose we begin with the alphabet {a, h, o, !}. Then we can consider the *laughing Santa vocabulary*, which consists of strings such as ‘hahaha!’, ‘hohoho!’, ‘hahahoho!’ ‘hohoha!’, and so on.<sup>70</sup> Here is a graphical representation of a *laughing Santa finite-state automaton*, which can read and write the laughing Santa vocabulary:

### The Laughing Santa Automaton



The numbered nodes represent the *states* of the automaton, and the alphabetically labeled arcs represent *state-transitions*. By convention, the starting state is represented by a square (State 1), and the final state by a circle with a thick border (State 4).

As a *reader* of the laughing Santa vocabulary, the task of this automaton is to process a finite string, and determine whether or not it is a licit string of the vocabulary. It processes the string one alphabetic character at a time, beginning in State 1. It recognizes the string if and only if (when and only when) it arrives at its final state, State 4. If the first character of the string is not an ‘h’, it remains stuck in State 1, and rejects the string. If the first character is an ‘h’, it moves to State 2, and processes the next character. If that character is not an ‘a’ or an ‘o’, it remains stuck in State 2, and rejects the string. If the character is an ‘a’ or an ‘o’, it moves to State 3. If the next character is an exclamation point, it moves to State 4, and recognizes the

<sup>70</sup> In the syntactic metalanguage for specifying vocabularies that I mentioned in the note above, this is the vocabulary (h|ho)\*!

string ‘ha!’ or ‘ho!’—the shortest ones in the laughing Santa vocabulary. If instead the next character is an ‘h’, it goes back to State 2, and repeats itself in loops of ‘ha’s and ‘ho’s any number of times until an exclamation point is finally reached, or it is fed a discordant character.

As a *writer* of the laughing Santa vocabulary, the task of the automaton is to produce only licit strings of that vocabulary, by a process that can produce any and all such strings. It begins in its initial state, State 1, and emits an ‘h’ (its only available move), changing to State 2. In this state, it can produce either an ‘a’ or an ‘o’—it selects one at random<sup>71</sup>—and goes into State 3. In this state, it can either tack on an exclamation point, and move into its final state, State 4, finishing the process, or emit another ‘h’ and return to State 2 to repeat the process. In any case, whenever it reaches State 4 and halts, the string it has constructed will be a member of the laughing Santa vocabulary.

I hope this brief rehearsal makes it clear how the constellation of nodes and arrows that makes up this directed graph represents the abilities to read and write (recognize and produce arbitrary strings of) the laughing Santa vocabulary.<sup>72</sup> What it represents is abilities that are *PV*-

---

71 As a matter of fact, it can be shown that every vocabulary readable/writeable by a non-deterministic finite-state automaton—such as the laughing Santa automaton—is also readable/writeable by a deterministic one. M. O. Rabin and D. Scott, "Finite Automata and their Decision Problems," *IBM Journal of Research and Development* 3, no. 2 (1959), pp. 115-25.

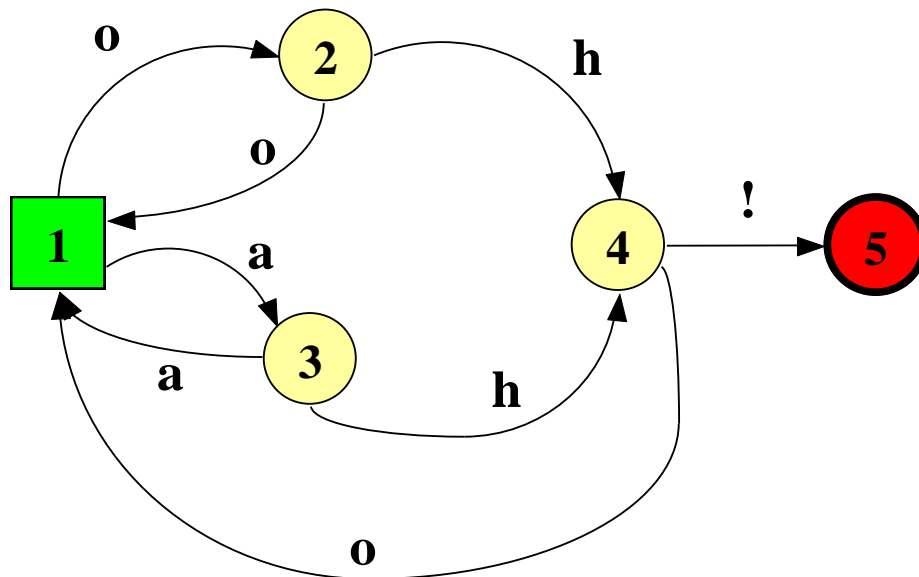
72 For practice, or to test one’s grip on the digraph specification of FSAs, consider what vocabulary over the same alphabet that produces the laughing Santa is recognized/produced by this automaton:

*sufficient to deploy* that vocabulary—that is, read and write it, in the attenuated sense appropriate to this purely syntactic case. And the digraph representation is itself a *vocabulary* that is *VP-sufficient to specify* those vocabulary-deploying abilities. That is, the digraph representation of this finite-state automaton is a *pragmatic metavocabulary* for the laughing Santa vocabulary. The relation between the digraph vocabulary and the laughing Santa vocabulary is, then, a *pragmatically mediated*—not now *semantic*, but *syntactic*—relation between vocabularies.

It may seem that I am stretching things by calling the digraph form of representation a ‘vocabulary’. It will be useful, as a way of introducing my final point in the vicinity, to consider a different form of pragmatic metavocabulary for the laughing Santa vocabulary. Besides the digraph representation of a finite-state automaton, we can also use a *state-table* representation. For the laughing Santa automaton this is:

---

### The "I'll Have What She's Having" Automaton



	<u>State 1</u>	<u>State 2</u>	<u>State 3</u>
<b>a</b>	Halt	3	Halt
<b>h</b>	2	Halt	2
<b>o</b>	Halt	3	Halt
<b>!</b>	Halt	Halt	4

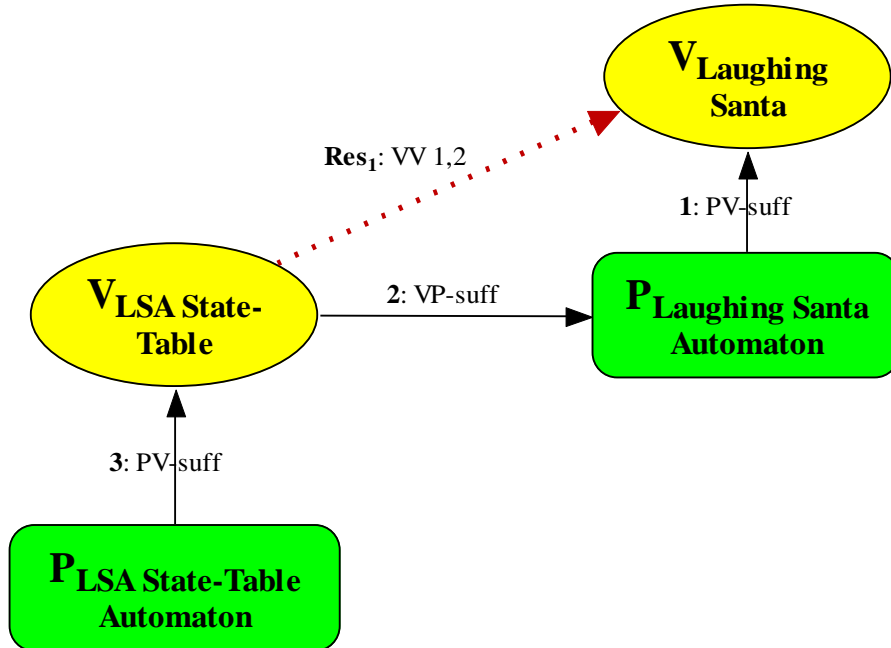
In read mode, the automaton starts in State 1. To see what it will do if fed a particular character, we look at the row labeled with that character. The LSA will Halt if the input string starts with anything other than an ‘h’, in which case it will change to State 2. In that state, the automaton specified by the table will halt unless the next character is an ‘a’ or an ‘o’, in which case it changes to State 3, and so on. (There is no column for State 4, since it is the final state, and accepts/produces no further characters.) Clearly there is a tabular representation corresponding to any digraph representation of an FSA, and vice versa. Notice further that we need not use a two-dimensional table to convey this information. We could put the rows one after another, in the form:

aHalt3Halth2Halt2oHalt3Halt!HaltHalt4.

This is just a string, drawn from a universe generated by the alphabet of the LSA, together with ‘Halt’ and the designations of the states of that automaton. The strings that specify finite-state automata that deploy vocabularies defined over the same basic alphabet as the LSA then form a vocabulary in the technical syntactic sense we have been considering. And that means we can ask about the automata that can read and write *those* state-table encoding vocabularies. The meaning-use diagram for this situation is then:



**Meaning-Use Diagram #6:  
Specifying the Automaton  
that Deploys the Laughing  
Santa Vocabulary**



## V. The Chomsky Hierarchy:

### A Syntactic Example of Pragmatic Expressive Bootstrapping

Restricting ourselves to a purely syntactic notion of a vocabulary yields a clear sense of ‘pragmatic metavocabulary’: both the digraph and the state-table vocabularies are VP-sufficient to specify practical abilities articulated as a finite-state automaton that is PV-sufficient to deploy—in the sense of recognizing and producing—the laughing Santa vocabulary, as well as many others. (Of course, it does that only against the background of a set of abilities PV-sufficient to deploy *those* vocabularies.) Perhaps surprisingly, it also offers a prime example of *strict pragmatic expressive bootstrapping*. For in this setting we can *prove* that one vocabulary that is expressively weaker than another can nonetheless serve as an adequate *pragmatic* metavocabulary for that stronger vocabulary. That is, even though one *cannot say* in the weaker vocabulary everything that can be *said* in the stronger one, one can still *say* in the weaker one everything that one needs to be able to *do* in order to deploy the stronger one.

Here the relevant notion of the relative expressive power of vocabularies is also a purely syntactic one. Already in the 1950’s, Chomsky offered mathematical characterizations of the different sets of strings of characters that could be generated by different classes of grammars (that is, in my terms, characterized by different kinds of syntactic metavocabularies) and computed by different kinds of automata. The kinds of vocabulary, grammar, and automata lined up with one another, and could be arranged in a strict expressive hierarchy: the Chomsky hierarchy. It is summarized in the following table:

<u>Vocabulary</u>	<u>Grammar</u>	<u>Automaton</u>
Regular	$A \rightarrow aB$ $A \rightarrow a$	Finite State Automaton
Context-Free	$A \rightarrow \langle \text{anything} \rangle$	Push-Down Automaton
Context-Sensitive	$c_1 A c_2 \rightarrow c_1 \langle \text{anything} \rangle c_2$	Linear Bounded Automaton
Recursively Enumerable	No Restrictions on Rules	Turing Machine (= 2 Stack PDA)

The point I want to make fortunately does not require us to delve very deeply into the information summarized in this table. A few basic points will suffice. The first thing to realize is that not all vocabularies in the syntactic sense we have been pursuing can be read and written by finite-state automata. For instance, it can be shown that no finite-state automaton is PV-sufficient to deploy the vocabulary  $a^n b^n$ , defined over the alphabet  $\{a, b\}$ , which consists of all strings of any arbitrary number of ‘a’s followed by the same number of ‘b’s. The idea behind the proof is that in order to tell whether the right number of ‘b’s follow the ‘a’s (when reading) or to produce the right number of ‘b’s (when writing), the automaton must somehow keep track of how many ‘a’s have been processed (read or written). The only way an FSA can store information is by being in one state rather than another. So, it could be in one state—or in one of a class of states—if one ‘a’ has been processed, another if two have, and so on. But by definition, a finite-state automaton only has a finite number of states, *and that number is fixed in advance* of receiving its input or producing its output. Whatever that number of states is, and whatever system it uses to code numbers into states (it need not be one-to-one—it could use a decimal coding, for instance), there will be some number of ‘a’s that is so large that the automaton runs out of states before it finishes counting. But the vocabulary in question consists of arbitrarily long strings of ‘a’s and ‘b’s. In fact, it is possible to say exactly which vocabularies

finite-state automata (specifiable by digraphs and state-tables of the sort illustrated above) *are* capable of deploying. These are called the ‘regular’ vocabularies (or languages).

The next point is that slightly more complex automata *are* capable of deploying vocabularies, such as  $a^n b^n$ , that are not regular, and hence cannot be read or written by finite-state automata. As our brief discussion indicated, intuitively the problem FSAs have with languages like  $a^n b^n$  is that they lack *memory*. If we give them a memory, we get a new class of machines: (non-deterministic<sup>73</sup>) *push-down automata* (PDAs). In addition to being able to respond differentially to and produce tokenings of the alphabetic types, and being able to change state, PDAs can *push* alphabetic values to the top of a *memory-stack*, and *pull* such values from the top of that stack. PDAs can do everything that finite-state automata can do, but they can also read and write many vocabularies, such as  $a^n b^n$ , that are *not* regular, and so cannot be read and written by FSAs. The vocabularies they can deploy are called “context free.” All regular vocabularies are context-free, but not *vice versa*. This proper containment of classes of vocabularies provides a clear sense, suitable to this purely syntactic setting, in which one vocabulary can be thought of as “expressively more powerful” than another: the different kinds of grammar can specify, and the different kinds of automata can compute, ever larger classes of vocabularies. Context-free vocabularies that are not regular require more powerful grammars to specify them, as well as more powerful automata to deploy them. FSAs are special kinds of PDAs, and all the automata are special kinds of Turing Machines. Recursively enumerable vocabularies are not in general syntactically reducible to context-sensitive, context-free, or regular ones. And the less capable automata cannot read and write all the vocabularies that can be read and written by Turing Machines.

Nonetheless, if we look at *pragmatically mediated* relations between these syntactically characterized vocabularies, we find that they make possible a kind of *strict expressive bootstrapping* that permits us in a certain sense to evade the restrictions on expressive power

---

<sup>73</sup> By contrast to FSA’s, there need not in general be, for every vocabulary computable by a non-deterministic PDA, some deterministic PDA that reads and writes the same vocabulary.

enforced for purely syntactic relations between vocabularies. The hierarchy dictates that only the abilities codified in Turing Machines—two-stack push-down automata—are *PV-sufficient* to *deploy* recursively enumerable vocabularies in general. But now we can ask: what class of languages is *VP-sufficient* to *specify* Turing Machines, and hence to serve as sufficient *pragmatic* metavocabularies for recursively enumerable vocabularies in general? The surprising fact is that **the abilities codified in Turing Machines—the abilities to recognize and produce arbitrary recursively enumerable vocabularies—can quite generally be specified in *context-free* vocabularies.** It is demonstrable that context-free vocabularies are strictly weaker in syntactic expressive resources than recursively enumerable vocabularies. The push-down automata that can read and write only context-free vocabularies cannot read and write recursively enumerable vocabularies in general. But it is possible to *say* in a context-free vocabulary what one needs to be able to *do* in order to deploy recursively enumerable vocabularies in general.

The proof of this claim is tedious, but not difficult, and the claim itself is not at all controversial—though computational linguists make nothing of it, having theoretical concerns very different from those that lead me to underline this fact. (My introductory textbook leaves the proof as an exercise to the reader.<sup>74</sup>) General-purpose computer languages such as Pascal and C++ can specify the algorithms a Turing Machine, or any other universal computer, uses to compute any recursively enumerable function, hence to recognize or produce any recursively enumerable vocabulary. And they are invariably context-free languages<sup>75</sup>—in no small part just because the simplicity of this type of grammar makes it easy to write parsers for them. Yet they suffice to specify the state-table, contents of the tape (or of the dual stacks), and primitive operations of any and every Turing Machine. Here is the MUD characterizing this pragmatically mediated relation between syntactically characterized vocabularies:

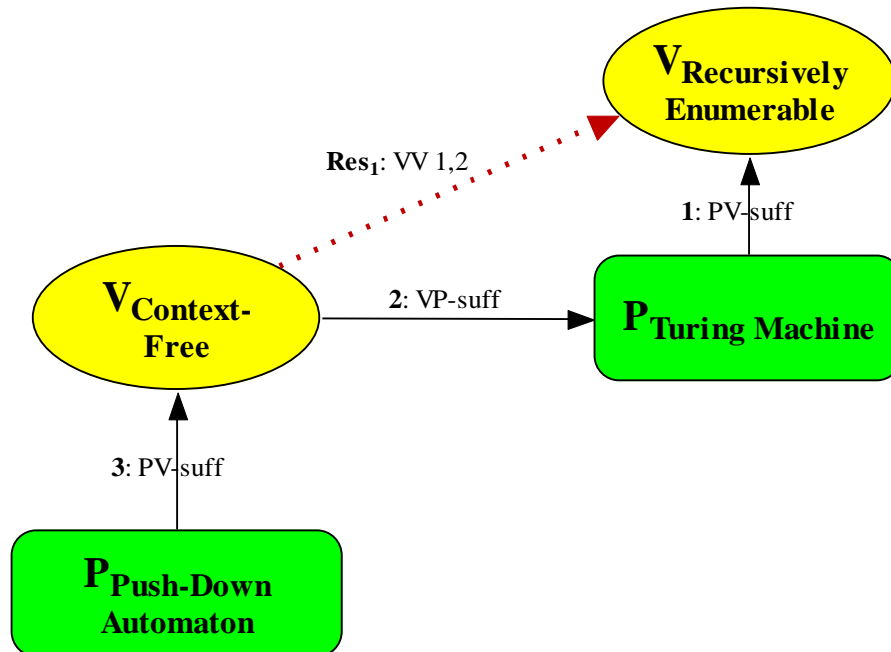
---

74 Thomas Sudkamp, *Languages and Machines*, 2nd ed. (Reading, MA: Addison-Wesley, 1997), Chapter 10.

75 In principle. There are subtleties that arise when we look at the details of actual implementations of particular computer languages, which can remove them from qualifying as strictly context-free.

### Meaning-Use Diagram #7:

#### Syntactic Pragmatic Expressive Bootstrapping



I called the fact that context-free vocabularies can be adequate pragmatic metavocabularies for recursively enumerable vocabularies in general ‘surprising’, because of the provable syntactic irreducibility of the one class of vocabularies to the other. But if we step back from the context provided by the Chomsky hierarchy, we can see why the possibility of such pragmatic expressive bootstrapping should not, in the end, be surprising. For all the result really means is that context-free vocabularies let one *say* what it is one must *do* in order to say things they cannot themselves say, because the ability to deploy those context-free vocabularies does not include the abilities those vocabularies let one specify. Thus, for instance, there is no reason that an FSA could not read and write a vocabulary that included commands such as “Push an ‘a’ onto the stack,”—and thus specify the program of a PDA—even though it itself has no stack, and could not *do* what the vocabulary it is deploying specifies. A coach might be able to tell an athlete exactly what to do, and even how to do it, even though the coach cannot himself do what he is telling the athlete to do, does not have the abilities he is specifying. We ought not to boggle at the possibility of an

expressively weaker pragmatic metavocabulary having the capacity to say what one must do in order to deploy an expressively stronger one. We should just look to see where this seems in fact to be possible for vocabularies we care about, and what we can learn from such relations when they do obtain.

## VI. Semantic Examples of Pragmatic Expressive Bootstrapping and Further Basic and Resultant Meaning-Use Relations

Let us recall what motivated this rehearsal of some elements of automaton theory and introductory computational linguistics. I suggested that a way to extend the classical project of semantic analysis so as to take account of the insights of its pragmatist critics is to look analytically at relations between meaning and use. More specifically, I suggested focusing to begin with on two in some sense complementary relations: the one that holds when some set of practices-or-abilities is PV-sufficient to *deploy* a given vocabulary, and the one that holds when some vocabulary is VP-sufficient to *specify* a given set of practices-or-abilities. The composition of these is the simplest *pragmatically mediated semantic relation* between vocabularies: the relation that holds when one vocabulary is a sufficient pragmatic metavocabulary for another. It is a paradigm of the infinite, recursively generable class of complex, pragmatically mediated semantic relations that I propose to lay alongside the other semantic relations between vocabularies that have been investigated by analytic philosophers (for instance those who address the core programs of empiricism and naturalism): relations such as analyzability, definition, translation, reduction, truth-making, and supervenience. I suggested further that pragmatic metavocabularies might be of particular interest in case they exhibited what I called “expressive bootstrapping”—cases, that is, in which the expressive power of the pragmatic metavocabulary differs markedly from that of the target vocabulary, most strikingly, when the metavocabulary is substantially expressively weaker—a phenomenon Tarski has led us *not* to expect for semantic metavocabularies, which in general must be expressively *stronger* than the vocabularies they address.

We have now seen that all of these notions can be illustrated with particular clarity for the special case of purely syntactically characterized vocabularies. The abilities that are PV-sufficient to deploy those vocabularies, in the sense of the capacity to recognize and produce them, can be thought of as various sorts of automata. There are several well-established, different-but-equivalent vocabularies that are known to be VP-sufficient to specify those



automata. In this special syntactic case we can accordingly investigate the properties of pragmatic metavocabularies, and when we do, we find a striking instance of *strict expressive bootstrapping* in a pragmatically mediated syntactic relation between vocabularies.

Of course, the cases we really care about involve *semantically* significant vocabularies. Are there any interesting instances of these phenomena in such cases? I have indicated briefly how some of Sellars's pragmatist criticisms of various ways of pursuing the empiricist program can be understood to turn on pragmatically mediated semantic relations. And I mentioned Huw Price's idea that although normative vocabulary is not semantically reducible to naturalistic vocabulary, naturalistic vocabulary might suffice to specify what one must *do*—the practices-or-abilities one must engage in or exercise—in order to deploy normative vocabulary. Here is another example that I want to point to, though I cannot develop the claim here. For roughly the first three-quarters of the twentieth century, philosophers who thought about indexical vocabulary took for granted some version of the doctrine that a tokening *n* of an expression of the type 'now' was *synonymous* with, definable or semantically analyzable as, 'the time of utterance of *n*,' and similarly for 'here' and 'the place of utterance of *h*,' and so on. During the 1970's philosophers such as John Perry, David Lewis, and G. E. M. Anscombe, by focusing on the use of indexicals in modal and epistemic contexts, showed decisively that this cannot be right: what is expressed by indexical vocabulary cannot be expressed equivalently by non-indexical vocabulary. This fact seems so obvious to us now that we might be led to wonder what philosophers such as Russell, Carnap, and Reichenbach could have been thinking for all those years. I want to suggest that the genuine phenomenon in the vicinity is a *pragmatically mediated* semantic relation between these vocabularies. Specifically, in spite of the *semantic* irreducibility of indexical to nonindexical vocabulary, it is possible to *say*, entirely in non-indexical terms, what one must *do* in order to be deploying indexical vocabulary correctly: to be saying essentially and irreducibly indexical things. For we can formulate practical rules such as:

1. If, at time *t* and place  $\langle x,y,z \rangle$ , speaker *s* wants to assert that some property *P* holds of  $\langle x,y,z,t,s \rangle$ , it is correct to say "P holds of *me*, *here* and *now*."

And

2. If a speaker  $s$  at time  $t$  and place  $\langle x, y, z \rangle$  asserts “ $P$  holds of *me, here and now*,” the speaker is committed to the property  $P$  holding of  $\langle x, y, z, t, s \rangle$ .

Non-indexical vocabulary can serve as an adequate *pragmatic* metavocabulary for indexical vocabulary. The fact that one nonetheless cannot *say* in non-indexical terms everything that one can *say* with indexical vocabulary just shows that these vocabularies have different expressive powers, so that the pragmatically mediated semantic relation between them is a case of strict pragmatic expressive bootstrapping.

Here is another example. Besides pragmatically mediated semantic relations between vocabularies, there is another sort of pragmatic analysis, which relates one constellation of practices-or-abilities to another. It corresponds to another basic meaning-use relation: the kind of PP-sufficiency that holds when having acquired one set of abilities means one can already do everything one needs to do, *in principle*, to be able to do something else. One concrete way of filling in a definite sense of “in principle” is by *algorithmic elaboration*, where exercising the target ability just is exercising the right basic abilities in the right order and under the right circumstances. (Of course, this is just *one* species of the genus of practical projection that Wittgenstein brings to our attention.) As an example, the ability to do long division just consists in exercising the abilities to do multiplication and subtraction according to a particular conditional branched-schedule algorithm. The practical abilities that *implement* such an algorithmic PP-sufficiency relation are just those exercised by finite-state *automata* (in general, Turing Machines). Indeed, automata should be thought of as consisting in a definite set of *meta-abilities*: abilities to *elaborate* a set of primitive abilities into a set of more complex ones, which can accordingly be pragmatically *analyzed* in terms of or *decomposed* into the other.<sup>76</sup>

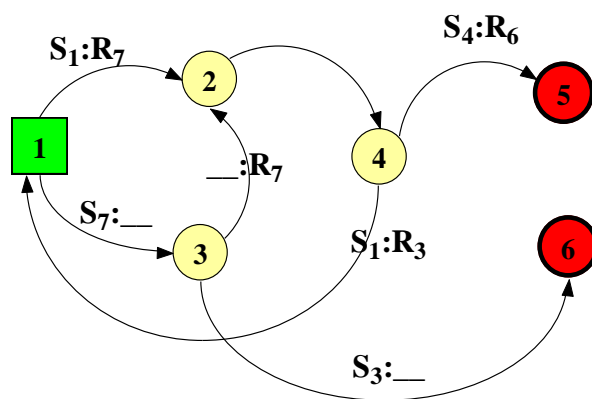
To get a usefully general concept of the PP-sufficiency of a set of basic abilities for a set of more complex ones, we need to move beyond the purely *syntactic* automata I have described so far.

---

<sup>76</sup> There are various vocabularies that are VP-sufficient for specifying those meta-abilities. Specifying them in terms of the differentially elicitable capacities to change state and to store and retrieve symbols is just one of them.

One way to do that is to replace their specialized capacities to read and write symbols—in the minimal sense of classifying tokens as to types and producing tokens of specified types—by more general recognitional and productive capacities. These are abilities to respond differentially to various in general non-symbolic stimuli (for instance, the visible presence of red things), corresponding to reading, and to respond by producing performances of various in general non-symbolic kinds (for instance, walking north for a mile), corresponding to writing. What practically implements the algorithmic elaboration of such a set of basic differential responsive abilities is a finite state *transducing* automaton (and its more sophisticated push-down brethren).

### A Finite-State Transducing Automaton



This is a diagram of an FSTA that has an initial set of stimuli to which it can respond differentially, and an initial set of responses it can differentially produce. And the diagram indicates that in its initial state, if presented with a stimulus of kind 1, it will produce a response of kind 7 and shift to state 2, and if presented instead with a stimulus of kind 7 it will produce no response, but will shift to state 3. It is important to note that although the recognitive and performative abilities that such an automaton algorithmically elaborates are to be considered as ‘primitive’ or ‘basic’ with respect to such elaboration, this does not mean that they are so in any absolute sense. The stimulus-response formulation by itself does not keep us from considering as ‘primitive’ capacities the abilities to keep ourselves at a suitable distance from a conversational partner, distinguish cubist paintings done by Braque from those done by Picasso, drive from New York to San Francisco, or build a house.

The notion of the algorithmic decomposability of some practices-or-abilities into others suggests in turn a pragmatic generalization of the classical program of artificial intelligence functionalism—which, though a latecomer in the twentieth century, deserves, I think, to count as a third core program of classical semantic analysis, alongside empiricism and naturalism. AI functionalism traditionally held itself hostage to a commitment to the purely *symbolic* character of intelligence in the sense of sapience. But broadening our concern from automata as purely syntactic engines to the realm of *transducing* automata, we are now in a position to see automaton functionalism as properly concerned with the algorithmic decomposability of discursive (that is, vocabulary-deploying) practices-and-abilities. What I will call the ‘pragmatic’ thesis of artificial intelligence is the claim that the ability to engage in some autonomous discursive practice (a language game one could play though one played no other) can be *algorithmically decomposed* into non-discursive abilities—where by “non-discursive” abilities, I mean abilities each of which can in principle be exhibited by something that does *not* engage in any autonomous discursive practice. (Without that restriction on the primitive abilities out of which discursive ones are to be algorithmically elaborated, the claim would be trivial, since the null algorithmic decomposition is also a decomposition.) The capacity to talk-and-think as I am addressing it is the capacity to deploy an autonomous vocabulary. But unlike classical symbolic AI, the pragmatic thesis of artificial intelligence does not presume that the practical capacities from which some transducing automaton can algorithmically elaborate the ability to engage in an autonomous discursive practice must themselves consist exclusively of symbol-manipulating abilities, never mind ultimately syntactic ones.<sup>77</sup>

The algorithmic practical elaboration model of AI gives a relatively precise shape to the pragmatist program of explaining knowing-*that* in terms of knowing-*how*: specifying in a non-intentional, non-semantic vocabulary what it is one must *do* in order to count as deploying some vocabulary to *say* something, hence as making intentional and semantic vocabulary applicable to

---

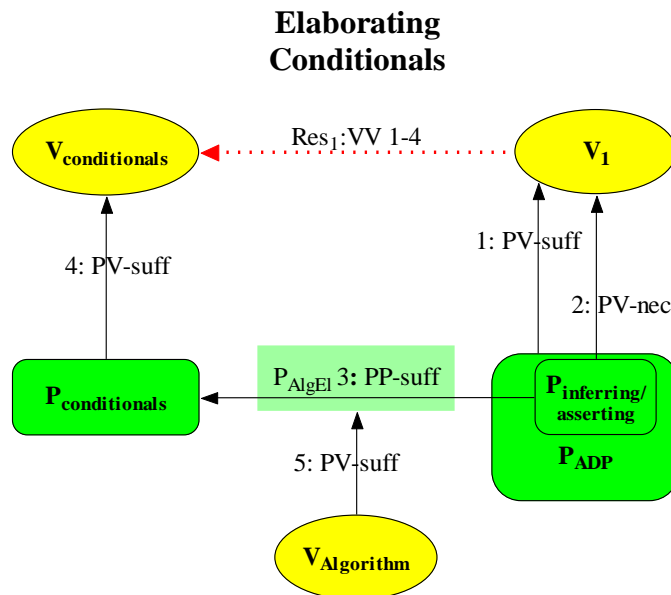
<sup>77</sup> For this reason, the frame problem, as it is often formulated, does not *immediately* arise for the pragmatic version of AI-functionalism. But as is explored in the third chapter of *Between Saying and Doing*, it does get a grip, at a different point.

the performances one produces. In particular, it offers a construal of the basic claim of AI-functionalism as a pragmatic *expressive bootstrapping* claim about computer languages as pragmatic metavocabularies for much more expressively powerful autonomous vocabularies, namely natural languages. The arguments for and against this pragmatic version of AI-functionalism accordingly look quite different from those arrayed on the opposing sides of the debate about the prospects of symbolic AI.

Combining the notion of PP-sufficiency that holds between two constellations of practices-or-abilities when one can be algorithmically elaborated from the other with the two sorts of basic meaning-use relations out of which I previously constructed the notion of expressively bootstrapping pragmatic metavocabularies—namely, a set of practices-or-abilities being PV-sufficient to *deploy* a vocabulary and a vocabulary being VP-sufficient to *specify* a set of practices-or-abilities—makes it possible to define further kinds of pragmatically mediated semantic relations. As my final example, consider the relation between *logical* vocabulary—paradigmatically, conditionals—and ordinary, non-logical, empirical descriptive vocabulary. I take it that every autonomous discursive practice must include performances that have the pragmatic significance of *assertions* and *inferences* (which I would argue come as an indissoluble package). I actually think this PP-necessary condition on any practices PV-sufficient for autonomously deploying a vocabulary can usefully be treated as sufficient as well—that is, as what distinguishes *discursive* practices as such. But nothing in what follows turns on that further commitment. To count as engaging in such practices, practitioners must exercise an ability, however fallible, to assess the goodness of material inferences: to sort them into those they accept and those they reject. This is part of what one must *do* in order to *say* anything. But it is easy to say how those recognitional and performative abilities, for these purposes counted as primitive, can be algorithmically elaborated into the capacity to use conditionals. An algorithm VP-sufficient to specify an automaton that practically implements such a pragmatic elaboration or PP-sufficiency relation is the following:

3. Assert the conditional ‘*if p then q*’ if one endorses the inference from *p* to *q*;
4. Endorse the inference from *p* to *q* if one asserts the conditional ‘*if p then q*’.

These rules of usage codify introduction and elimination rules for the conditional. So the capacity to use conditionals can be algorithmically elaborated from the capacities to make assertions and assess inferences. This is the composition of a PP-sufficiency relation with a PV-sufficiency relation, and is expressed in the following meaning-use diagram<sup>78</sup>:

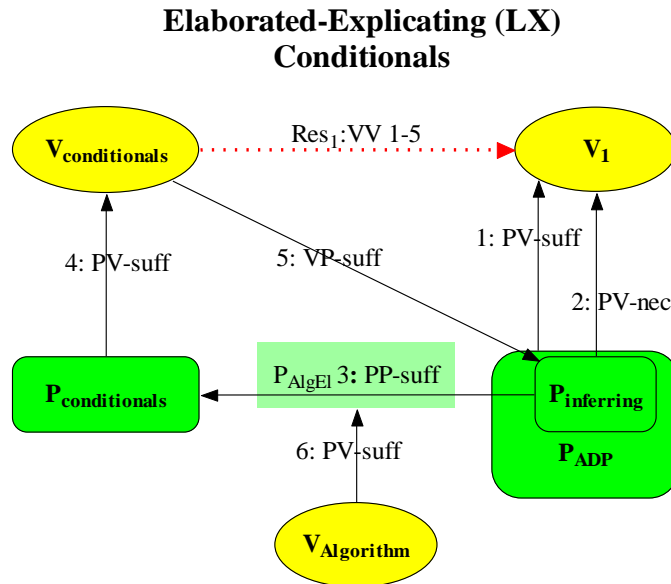


The complex resultant meaning-use relation indicated by the dotted arrow at the top of the diagram is a further pragmatically mediated semantic relation. The diagram indicates *exactly* what constellation of sub-claims about *basic* meaning-use relations must be justified in order to justify the claim that this relation obtains between two vocabularies, and hence the diagram graphically presents a distinctive kind of meaning-use *analysis* of that semantic relation.

In fact, if we think further about this example, by filling in another basic meaning-use relation that obtains in this case, we can define an even more articulated pragmatically mediated semantic relation between vocabularies. For when conditionals are deployed with the practical circumstances and consequences of application specified in the algorithm stated above, they let practitioners *say* what otherwise they could only *do*; that is, they express *explicitly*, in the form of

<sup>78</sup> I indicate PP-necessity relations by including the rounded rectangle for one set of practices-or-abilities in another.

a claimable, hence propositional, content, what practitioners are *implicitly* doing in endorsing some material inferences and rejecting others. This is a VP-sufficiency relation: conditionals let one *specify* the practices of taking-or-treating inferences as materially good or bad. Adding in this *explicating* relation between conditionals and the practices-or-abilities they make explicit yields a new pragmatically mediated semantic relation that conditionals stand in to every autonomously deployable vocabulary. Its meaning-use diagram is this:



The practical capacity to deploy conditionals (that is, something PV-sufficient for their use) both can be *elaborated from* practices PP-*necessary* for every ADP, and *explicates* those practices (in the sense of being VP-sufficient for them). It is *elaborated-explicative* relative to every autonomous vocabulary. We say, it is LX for every AV, hence for every vocabulary (since the use of any vocabulary presupposes, and in that sense is parasitic on, the capacity to use some autonomous vocabulary).

I believe that this complex resultant pragmatically mediated semantic relation is important for understanding the distinctive semantic role played by *logical* vocabulary generally: not just conditionals, but also negation (which makes explicit a central feature of our practice of treating claims as materially *incompatible*), and even modal vocabulary (which makes explicit a central feature of our practice of associating *ranges of counterfactual robustness* with material

inferences). In my initial characterization of the classical semantic project of philosophical analysis, I pointed to the special status that is accorded to logical vocabulary in that project. What I called “semantic logicism” is its commitment to the legitimacy of the strategy of using logical vocabulary to articulate the semantic relations between vocabularies that is its goal—paradigmatically in connection with the core projects of empiricism, naturalism, and functionalism. One interesting way to vindicate that commitment (that is, at once to explain and to justify it) would be to appeal to the fact that logical vocabulary is elaborated from and explicating of every autonomously deployable vocabulary whatsoever. For that means that the capacity to use logical vocabulary is both in this very clear and specific sense *implicit* in the capacity to use any vocabulary, and has the expressive function of making *explicit* something already present in the use of any vocabulary.

I won't say anything more here about how such a vindication might proceed, contenting myself with the observation that insofar as there is anything to an account along these lines, supplementing the traditional philosophical analytical concern with semantic relations between the *meanings* expressed by different kinds of vocabulary by worrying also about the pragmatic relations between those meanings and the *use* of those vocabularies in virtue of which they express those meanings is not so much *extending* the classical project of analysis as *unpacking* it, to reveal *explicitly* a pragmatic structure that turns out to have been *implicit* in the analytic semantic project all along. For the conclusion will be that it is *because* some vocabularies are universal pragmatically elaborated and explicating vocabularies that semantic analysis of the logicist sort is both possible and legitimate at all. I don't claim to have entitled myself to that conclusion here, only to have introduced some conceptual machinery that might make it possible to do so—and so at least to have sketched a way in which the insights of the *pragmatist* tradition can be assembled and developed so as to be constructively helpful to, rather than destructively critical of, the classical project of philosophical semantic *analysis*, and so to open the way to extending that project in promising new directions.



## **2014 Nordic Pragmatism Lectures:**

### **Analytic Pragmatism, Expressivism, and Modality**

#### **Lecture 4**

#### **Pragmatism, Inferentialism, and Modality in Sellars's Arguments against Empiricism**

##### **I. Introduction**

In this lecture I want to place the arguments of “Empiricism and the Philosophy of Mind” into the context of some of Sellars’s other, nearly contemporary articles, by tracing further, into those neighboring works, some strands of argumentation that intersect and are woven together in his critique of empiricism in its two principal then-extant forms: traditional, and twentieth-century logical empiricism. Sellars always accepted that observation reports resulting non-inferentially from the exercise of perceptual language-entry capacities play both the privileged epistemological role of being the ultimate court of appeal for the justification of empirical knowledge-claims and therefore (given his inferentialist semantics) an essential semantic role in determining the contents of the empirical concepts applied in such judgments. But in accord with his stated aspiration to “move analytic philosophy from its Humean into its Kantian phase,” he was severely and in principle critical of empiricist ambitions and programs in epistemology and (especially) semantics that go beyond this minimal, carefully circumscribed characterization of the cognitive significance of sense experience. Indeed, I think the lasting philosophical interest of Sellars’s thought lies primarily in the battery of original considerations and arguments he brings to bear against all weightier forms of empiricism. Some, but not all, of these are deployed in the opening critical portions of “Empiricism and the Philosophy of Mind,” where the

ground is cleared and prepared for the constructive theorizing of the last half. But what is on offer there is only part of Sellars's overall critique of empiricism. We accordingly court misunderstanding of what is there if we do not appreciate the shape of the larger enterprise to which it contributes.

In an autobiographical sketch, Sellars dates his break with traditional empiricism to his Oxford days in the thirties. It was, he says, prompted by concern with understanding the sort of conceptual content that ought to be associated with "logical, causal, and deontological modalities." Already at that point he says that he had the idea that

what was needed was a functional theory of concepts which would make their role in reasoning, rather than supposed origin in experience, their primary feature.<sup>79</sup>

This telling passage introduces two of the master ideas that shape Sellars's critique of empiricism. The first is that a key criterion of adequacy with respect to which its semantics will be found wanting concerns its treatment of *modal* concepts. The second is that the remedy for this inadequacy lies in an alternative broadly functional approach to the semantics of these concepts that focuses on their *inferential roles*—as it were, looking *downstream* to their subsequent *use*, as well as *upstream* to the circumstances that elicit their application.

This second, inferential-functionalist, semantic idea looms large in "Empiricism and the Philosophy of Mind." In fact, it provides the raw materials that are assembled and articulated into Sellars's positive account of the semantics of the concepts applied in reporting thoughts and sense-impressions. Concern with the significance of *modality* in the critique of empiricism, however, is almost wholly absent from that work (even though it is evident in articles Sellars wrote even earlier). I do not think that is because it was not, even then, an essential element of the larger picture of empiricism's failings that Sellars was seeking to convey, but rather because

---

<sup>79</sup> "Autobiographical Reflections (February, 1973)," in *Action, Knowledge, and Reality*, ed. H. N. Castañeda (Indianapolis: Bobbs-Merrill, 1975), p. 285.

it was the result of a hard-won but ultimately successful divide-and-conquer expository strategy. That is, I conjecture that what made it possible for Sellars finally to write “Empiricism and the Philosophy of Mind” was figuring out a way to articulate the considerations he advances there *without* having also at the same time to explore the issues raised by empiricism’s difficulties with modal concepts. Whether or not that conjecture about the intellectual-biographical significance of finding a narrative path that makes possible the separation of these aspects of his project is correct, I want to claim that it is important to understand what goes on in “Empiricism and the Philosophy of Mind” in the light of the fuller picture of the expressive impoverishment of empiricism that becomes visible when we consider what Sellars says when he *does* turn his attention to the semantics of modality.<sup>80</sup>

There is a third strand to the rope with which Sellars first binds and then strangles the excessive ambitions of empiricism. That is his methodological strategy of considering *semantic* relations among the meanings expressed by different sort of vocabulary that result from *pragmatic* dependencies relating the practices one must engage in or the abilities one must exercise in order to count as using those bits of vocabulary to express those meanings. This is the *pragmatist* element in Sellars’s multi-front assault on empiricism. It makes a significant contribution to the early, critical portion of EPM, though Sellars does not overtly mark it, as he does the contribution of his inferential functionalism to the later, more constructive portion. The concern with what one must *do* in order to *say* (so, to think) various kinds of things remains implicit in what Sellars *does*, rather than explicit in what he *says about* what he *does*. As we will see, both the pragmatist and the inferentialist ideas are integral to his critique of empiricist approaches to modality and to his constructive suggestions for a more adequate treatment of modal vocabulary.

---

<sup>80</sup> *Empiricism and the Philosophy of Mind* reprinted with an introduction by Richard Rorty and a study guide by Robert Brandom (Cambridge, MA: Harvard University Press, 1997), hereafter cited as EPM.

## II. The Inferentialist and Pragmatist Critique of Empiricism in *EPM*

I think of the classical project of analytic philosophy in the twentieth century as being the exploration of how the meanings expressed by some target vocabularies can be exhibited as in some sense a logical elaboration of the meanings already expressed by some base vocabularies. The conception of the desired semantic relation between vocabularies (the sense of ‘analysis’) varied significantly within this broadly defined semantic project, including definition, paraphrase, translation, reduction in various senses, supervenience, and truth-making, to name just a few prominent candidates. I take it to be integral to the analytic philosophical project during this period that however that semantic relation is conceived, *logical* vocabulary is taken to play a special role in elaborating the base vocabulary into the target vocabulary. The distinctively twentieth-century form of *empiricism* can be understood as one of the *core programs* of this analytic project—not in the sense that every participant in the project endorsed some version of empiricism (Neurath, for instance, rejects empiricism where he sees it clashing with another core semantic program that was dearer to his heart, namely naturalism), but in the sense that even those who rejected it for some target vocabulary or other took the possibility of an empiricist analysis to be an important issue, to set a legitimate philosophical agenda.

Construed in these terms, twentieth-century empiricism can be thought of as having proposed three broad kinds of empiricist base vocabularies. The most restrictive kind comprises *phenomenalist* vocabularies: those that specify how things subjectively *appear* as opposed to how they objectively are, or the not-yet-conceptualized perceptual *experiences* subjects have, or the so-far-uninterpreted sensory *given* (the data of sensation: sense data). A somewhat less restrictive genus of empiricist base vocabularies limits them to those that express *secondary qualities*, thought of as what is *directly perceived* in some less demanding sense. And a still more relaxed version of empiricism restricts its base vocabulary to the *observational* vocabulary deployed in non-inferentially elicited perceptual reports of observable states of affairs. Typical target vocabularies for the first, phenomenalist, class of empiricist base vocabularies include

those expressing empirical claims about how things really or *objectively* are—that is, those expressing the applicability of any objective empirical concepts. Typical target vocabularies for secondary-quality empiricism include any that specify *primary qualities* or the applicability of concepts that are not response-dependent. And typical target vocabularies for observational vocabulary empiricism include *theoretical* vocabulary. All species of empiricism are concerned with the possibility of underwriting a semantics for the modal vocabulary used to express laws of nature, probabilistic vocabulary, normative vocabulary, and others sophisticated vocabularies of independent philosophical interest. The standard empiricist alternatives are either to show how a given target vocabulary can be semantically elaborated from the favored empiricist base vocabulary, on the one hand, or to show how to live with a local skepticism about its ultimate semantic intelligibility, on the other.

At the center of Sellars's critique of empiricism in EPM is an argument against the weakest, least committive, observational, version of empiricism (a critique that then carries over, *mutatis mutandis*, to the more demanding versions). That argument depends on both his inferential-functionalist semantics and on his pragmatism. Its fundamental strategy is to show that the proposed empiricist base vocabulary is not *pragmatically* autonomous, and hence not *semantically* autonomous. Observational vocabulary is not a vocabulary one could use though one used no other. Non-inferential reports of the results of observation do not form an autonomous stratum of language. In particular, when we look at what one must *do* to count as making a non-inferential report, we see that that is not a practice one could engage in except in the context of *inferential* practices of using those observations as *premises* from which to draw inferential *conclusions*, as *reasons* for making judgments and undertaking commitments that are *not* themselves observations. The contribution to this argument of Sellars's inferential functionalism about semantics lies in underwriting the claim that for *any* judgment, claim, or belief to be contentful in the way required for it to be *cognitively*, *conceptually*, or *epistemically* significant, for it to be a potential bit of *knowledge* or *evidence*, to be a *sapient* state or status, it must be able to play a distinctive *role in reasoning*: it must be able to serve as a *reason for* further judgments, claims, or beliefs, hence as a *premise* from which they can be *inferred*. That role in reasoning, in particular, what those judgments, claims, or beliefs can serve as reasons or

evidence *for*, is an essential, and not just an accidental component of their having the semantic content that they do. And that means that one cannot count as understanding, grasping, or applying concepts *non-inferentially* in observation unless one can also deploy them at least as premises in *inferences* to conclusions that do *not*, for that very reason, count as *non-inferential* applications of concepts. Nor, for the same reason, can any discursive practice consist entirely of non-inferentially acquiring *premises*, without any corresponding practice of drawing *conclusions*. So non-inferential, observational uses of concepts do not constitute an autonomous discursive practice: a language game one could play though one played no other. And this conclusion about the pragmatic dependence of observational uses of vocabulary on inferential ones holds no matter what the subject-matter of those observations is: whether it is observable features of the external environment, how things merely appear to a subject, or the current contents of one's own mind.

Here the pragmatist concern with what one must *do* in order to be able to *say* (or think) something combines with semantic inferentialist-functionalism about conceptual content to argue that the proposed empiricist base vocabulary is not pragmatically autonomous—since one must be able to make claims inferentially in order to count as making any non-inferentially. If that is so, then potentially risky inferential moves cannot be seen as an in-principle optional superstructure erected on a semantically autonomous base of things directly known through observation.

Although this is his most general and most powerful argument, Sellars does not limit himself to it in arguing against the substantially more committive forms of empiricism that insist on phenomenalist base vocabularies. In addition, he develops a constructive account of the relations between (at least one principle species of) phenomenalist vocabulary and objective vocabulary that depends on pragmatic dependences between what one must *do* in order to deploy each kind, to argue once again that the proposed empiricist base vocabulary does not form a semantically autonomous stratum of the language. This is his account of the relation between 'looks'-talk and 'is'-talk.

It develops out of his positive account of what one must *do* in order to use vocabulary observationally. To apply the concept green non-inferentially one must be able to do at least two sorts of things. First, one must be able reliably to respond differentially to the visible presence of green things. This is what blind and color-blind language-users lack, but non-language-using pigeons and parrots possess. Second, one must be able to exercise that capacity by reliably responding differentially to the visible presence of green things by applying the *concept green*. So one must possess, grasp, or understand that concept. “Grasp of a concept is mastery of the use of a word,” Sellars says, and his inferential functionalism dictates that this must include the *inferential* use of the word: knowing at least something about what follows from and is evidence for or against something’s being green. This the blind or color-blind language-user has, and the pigeon and parrot do not. Only the performances of the former can have the pragmatic significance of taking up a stand in the space of reasons, of committing themselves to something that has a *conceptual*, that is, inferentially articulated, content.

The point of Sellars’s parable of John in the tie shop is to persuade us that the home language game of the ‘looks’ or ‘seems’ vocabulary that expresses how things merely appear to us, without undertaking any commitment to how they actually are, is one that is pragmatically parasitic on the practice of making in-principle risky reports of how things objectively are. For what one must *do* in order to count as saying how things merely *look*, Sellars claims, is to *evince* the reliable differential disposition to respond to something by claiming that it is green, while *withholding* the endorsement of that claim (because of one’s collateral beliefs about the situation and one’s reliability in it). If that is what one is doing in making a ‘looks’-claim, then one cannot be wrong about it in the same way one can about an ‘is’-claim, because one has withheld the principal commitment rather than undertaking it. And it follows that phenomenalist ‘looks’-talk, which expresses how things merely appear, without further commitment to how things actually are, is not an autonomous discursive practice—not a language game one could play though one played no other—but is in fact pragmatically parasitic on objective ‘is’-talk.

My point in rehearsing this familiar argument is to emphasize the role played both by Sellars's pragmatist emphasis on what one must be able to *do* in order count as *saying* various kinds of thing—*using* vocabulary so as to express certain kinds of meanings—and by his inferentialist-functionalist insistence that the role some vocabulary plays in *reasoning* makes an essential contribution to its semantic content. Although Sellars does not go on to make this argument, the way these two lines of thought conspire to undermine the semantic autonomy of candidate empiricist base vocabularies provides a template for a parallel objection to secondary-quality empiricism. For at least a necessary condition on anything's being a secondary-quality concept is that it have an observational role that supports the introduction of corresponding 'looks'-talk, so that mastery of that 'looks'-talk can be taken to be essential to mastery of the concept—as 'looks-green' arguably is for mastery of the concept green, but 'looks'-square is *not* for mastery of the concept square. What would be needed to fill in the argument against secondary-quality empiricism via the non-autonomy of its proposed base vocabulary, would be an argument that nothing could count as mastering a vocabulary consisting entirely of expressions of this sort, apart from all inferential connections to primary-quality concepts that did not have this structure.



### III. Pragmatism and Phenomenalism

Thus far I have confined myself to offering a general characterization of anti-empiricist arguments that appear in “Empiricism and the Philosophy of Mind.” None of them involve empiricism’s treatment of modality. Now I want to put those arguments in a somewhat different frame, by conjoining them with one that is presented elsewhere, and which *does* turn on the significance of modal concepts. The previous arguments concerned the suitability of some vocabulary to serve as the *base* vocabulary of an empiricist analysis—since plausible motivations for caring about such an analysis typically require that it be semantically autonomous. This one turns on the criteria of adequacy of the analysis itself. My remarks in this section concern Sellars’s arguments in his essay “Phenomenalism,” which can be regarded as a kind of companion piece to EPM. (Later I will discuss another contemporary essay that I think should be thought of as yoked together with these two in a troika.) The first, modal, point is one that Sellars registers there, but does not linger on—his principal concern being rather with a second point, concerning another aspect of the vocabulary in which phenomenalist analyses would have to be couched. But given my purposes here, I want to make a bit more of the modal point than he does.

The basic idea of a phenomenalist-empiricist semantic analysis of ordinary objective vocabulary is that the expressive work done by talk of mind-independent objects and their properties and relations can be done by talk of *patterns* in, *regularities* of, or *generalizations* concerning sense experiences characterized in a phenomenalist vocabulary. Saying that the curved red surface I am experiencing is an experience *of* an apple that has parts I am *not* experiencing—a similarly bulgy, red back and a white interior, for instance—is properly understood as saying something about what I *would* experience if I turned it around or cut it open. That it continued to exist in the kitchen when I left the room is a matter of what I *would* have experienced *had* I returned. The first, obvious, observation is that an account of objective reality in terms of the *powers* of circumstances to *produce*, or my *dispositions* to *have*,

sensations, experiences, beings-appeared-to and so on essentially involves *modal* concepts. The patterns, regularities, or generalizations in subjective appearances that are supposed to constitute objective realities are modally robust, counterfactual-supporting patterns, regularities, or generalizations. Talk of what I actually *do* experience will not by itself underwrite claims about unexperienced spatial or temporal parts of empirical objects. Twentieth-century logical empiricism promised to advance beyond traditional empiricism because it could call on the full expressive resources of *logical* vocabulary to use as the ‘glue’ sticking sensory experiences together so as to construct simulacra of external objects. But *extensional* logical vocabulary is not nearly expressively powerful enough for the phenomenalist version of the empiricist project. So the phenomenalist conditional “terminating judgments,” into an infinite set of which C. I. Lewis proposes (in his *Analysis of Knowledge and Valuation*) to translate the “non-terminating judgments” of ordinary objective empirical discourse, have to use his modal notion of *strict* or necessary implication.<sup>81</sup> And similar points could be made about other phenomenalist reductionists such as Ayer. The consequence of this observation to which I want to draw attention is that one cannot use such a strategy in one’s phenomenalist empiricist analysis, translation, or reduction of objective talk *and* at the same time be a Humean skeptic about what *modal* vocabulary expresses. So essential features of the only remotely plausible *constructive* strategy of phenomenalist empiricism are simply incompatible with the most prominent *skeptical* consequences about modal concepts characteristically drawn both by traditional and twentieth-century logicist empiricism.

This is a powerful argument. Sellars’s principal concern in his essay “Phenomenalism,” however, is with a subsequent point. The conditionals codifying the patterns, regularities, or generalizations concerning sense experience that correspond to judgments about how things objectively are must not only be subjunctive, counterfactually robust conditionals, but in order to have any hope of being materially adequate (getting the truth-conditions even approximately correct) their *antecedents* must themselves be expressed in *objective* vocabulary, *not* in *phenomenalist* vocabulary. What is true (enough) is that if I were *actually* to turn the apple

---

<sup>81</sup> La Salle, IL: Open Court, 1946.

around, cut it open, or return to its vicinity in the kitchen I *would* have certain sense experiences. It is *not* in general true that if I merely *seem* to do those things I am guaranteed to have the corresponding experiences. For, phrased in such phenomenalist terms, the antecedent is satisfied in cases of imagination, visual illusion, dreaming, hallucination and so on that are precisely those *not* bound by the supposedly object-constituting rules and regularities. As Sellars summarizes the point:

To claim that the relationship between the framework of sense contents and that of physical objects can be construed on the [phenomenalist] model is to commit oneself to the idea that there are inductively confirmable generalizations about sense contents which are 'in principle' capable of being formulated without the use of the language of physical things.... [T]his idea is a mistake.<sup>82</sup>

It is a mistake because:

[T]he very selection of the complex patterns of actual sense contents in our past experiences which are to serve as the antecedents of the generalizations in question presuppose our common sense knowledge of ourselves as perceivers, of the specific physical environment in which we do our perceiving and of the general principles which correlate the occurrence of sensations with bodily and environmental conditions. We select those patterns which go with our being in a certain perceptual relation to a particular object of a certain quality, where we know that being in this relation to an object of that quality normally eventuates in our having the sense content referred to in the consequent.

This argument then makes evident

the logical dependence of the framework of private sense contents on the public, inter-subjective, logical space of persons and physical things.<sup>83</sup>

---

<sup>82</sup> "Phenomenalism" in *In the Space of Reasons*, ed. Kevin Scharp and Robert Brandom [Cambridge, MA: Harvard University Press, 2007], p. 331.

<sup>83</sup> "Phenomenalism," p. 328.

So the phenomenalist vocabulary is not autonomous. It is not a language game one can play though one plays no other. In particular, the uses of it that might plausibly fulfill many of the same pragmatic functions as ordinary objective empirical talk themselves presuppose the ability to deploy such objective vocabulary.

As Sellars points out, the lessons learned from pressing on the phenomenalist version of empiricism apply more generally. In particular, they apply to the more liberal version of empiricism whose base vocabulary is observational, including observations of enduring empirical objects, and whose target vocabulary is theoretical vocabulary. To begin with, if talk of theoretical entities is to be translated into or replaced by talk of patterns in, regularities of, or generalizations about observable entities, they must be *lawlike*, *counterfactual*-supporting regularities and generalizations. They must permit inferences to what one *would* observe if one *were* to find oneself in specified circumstances, or to prepare the apparatus in a certain way. For, once again, the patterns, regularities, or generalizations about observations the assertion of which an instrumentalist empiricist might with some initial plausibility take to have the same pragmatic effect as (to be doing the same thing one is doing in) deploying theoretical vocabulary must reach beyond the parochial, merely autobiographically significant contingencies of what subjects happen actually to observe. The theory is that electrical currents cause magnetic fields regardless of the presence of suitable measuring devices. And that can only be made out in terms of what is *observable*, that is *could* be observed, not just what *is* observed. And that is to say that the instrumentalist-observational form of empiricism is also incompatible with Humean-Quinean skepticism about the intelligibility of what is expressed by alethic modal vocabulary.

And an analog of the second argument against phenomenalist forms of empiricism also applies to instrumentalist forms. For, once again, the *antecedents* of the counterfactual conditionals specifying what *could* or *would* have been observed *if* certain conditions *had* obtained or certain operations *were* performed cannot themselves be formulated in purely observational terms. The meter-needle *would* have been observably displaced if I had connected the terminals of a volt-ohmmeter to the wire, but that something *is* a VOM is *not* itself a fact restatable in purely observational terms. Even leaving apart the fact that it is a *functional*

characterization not equivalent to any specification in purely *physical* terms, a description of the construction of some particular kind of VOM is still going to help itself to notions such as being made of copper, or being an electrical insulator (another bit of vocabulary that is both functional and theoretical). To satisfy the semantic ambitions of the instrumentalist it is not enough to associate each theoretical claim with a set of jointly pragmatically equivalent counterfactual-supporting conditionals whose *consequents* are couched wholly in observational vocabulary. All the theoretical terms appearing in the *antecedents* of those conditionals must be similarly replaced. No instrumentalist reduction of any actual theoretical claim has ever been suggested that even attempts to satisfy this condition.

Though Sellars does not, and I will not, pursue the matter, one expects that corresponding arguments will go through, *mutatis mutandis*, also for the kind of empiricism that seeks to understand the use of primary-quality vocabulary wholly in terms of the use of secondary-quality vocabulary. What we mean by talk of primary qualities will have to be cashed out in terms of its *powers* to produce, or our *dispositions* to perceive, secondary qualities—that is, in terms of modally robust, counterfactual-supporting generalizations. And it will be a challenge to specify the antecedents of a materially adequate set of such conditionals wholly in the official secondary-quality vocabulary.

#### IV. Sellars's Pragmatism and Modality

The arguments I have considered so far set limits to the semantic ambitions of phenomenalist and instrumentalist forms of analytic empiricism, first by focusing on the *pragmatic* preconditions of the required semantic autonomy of the proposed empiricist base vocabularies, and second by looking in more detail at the specific sorts of *inferential* patterns in the base vocabulary in terms of which it is proposed to reconstruct the circumstances and consequences of application of items in the various target vocabularies. Here it was observed that the material adequacy of such reconstructions seems to require the ineliminable involvement of terms from the target vocabulary, not only on the right side, but also on the left side of any such reconstruction—in the *definiens* as well as in the *definiendum*. Modality plays a role in these arguments only because the material adequacy of the reconstruction also turns out to require appeal to counterfactually robust inferences in the base vocabulary. Insofar as that is so, the *constructive* semantic projects of the phenomenalist, instrumentalist, and secondary-quality forms of empiricism are at odds with the local semantic skepticism about what is expressed by alethic modal vocabulary that has always been a characteristic cardinal *critical* consequence of empiricist approaches to semantics, as epitomized for its traditional phase by Hume and for its logicist phase by Quine.

In another massive, pathbreaking essay of this period, “Counterfactuals, Dispositions, and the Causal Modalities”<sup>84</sup> (completed in February of 1957), Sellars argues directly against this empiricist treatment of modality, completing what then becomes visible as a two-pronged attack on the principal contentions and projects of empiricism, only the opening salvos of which were fired in “Empiricism and the Philosophy of Mind.”<sup>85</sup> His principal target here is the “tendency to

---

<sup>84</sup> In *Minnesota Studies in the Philosophy of Science*, vol. 2, ed. H Feigl, M. Scriven, and G. Maxwell (Minneapolis: University of Minnesota Press, 1957), pp. 225-308. Hereafter CDCM.

<sup>85</sup> As in EPM (and even, though to a lesser extent, in “Phenomenalism”), in this essay Sellars describes himself not as denying empiricism, but rather as correcting it, protecting its core insights from the damage done by their over-

assimilate all discourse to describing,” which he takes to be primarily “responsible for the prevalence in the empiricist tradition of ‘nothing-but-ism’ in its various forms (emotivism, philosophical behaviorism, phenomenism).”<sup>86</sup> The form Sellars addresses in this essay is the Humean one that can find in statements of laws of nature, expressed in alethic modal vocabulary that lets us say what is and is not necessary and possible, “nothing but” expressions of matter-of-factual regularities or constant conjunctions (though he claims explicitly that considerations corresponding to those he raises for causal modalities are intended to apply to logical and deontological modalities as well<sup>87</sup>). His arguments are directed against the view that holds modal vocabulary semantically unintelligible, on grounds of inability to specify what it is saying about what the world is like, how it is describing things as being, insofar as by using it we are asserting something that goes beyond endorsing the existence of non-modally characterizable universal descriptive generalizations.

Hume found that even his best understanding of actual observable empirical *facts* did not yield an understanding of *rules* relating or otherwise governing them. Those facts did not settle which of the things that *actually* happened *had* to happen (given others), that is, were (at least conditionally) *necessary*, and which of the things that did *not* happen nonetheless were *possible* (not ruled out by laws concerning what did happen). The issue here concerns the justifiability and intelligibility of a certain kind of *inference*: modally robust, counterfactual-supporting inferences, of the kind made explicit by the use of modal vocabulary. Hume (and, following him, Quine) took it that epistemologically and semantically fastidious philosophers face a stark choice: either show how to explain modal vocabulary—the circumstances of application that justify the distinctive counterfactual-supporting inferential consequences of application—in

---

extension. But he also makes it clear that the result of such rectification is a Kantian view that gives equal weight to rationalist insights, when they are suitably reconstructed. So for instance he says:

It is my purpose to argue that the core truth of Hume’s philosophy of causation is not only compatible with, but absurd without, *ungrudging* recognition of those features of causal discourse as a mode of rational discourse on which the ‘metaphysical rationalists’ laid such stress, but also mis-assimilated to describing.” (CDCM §82)

And the final sentence of the essay invokes the “profound truth” of Kant’s conception of reason, “which empiricism has tended to distort.”

<sup>86</sup> CDCM §103.

<sup>87</sup> *Ibid.*

nonmodal terms, or show how to live without it, to do what we need to do in science without making such arcane and occult supradescriptive commitments.

This demand was always the greatest source of tension between empiricism and naturalism, especially the scientific naturalism that Sellars epitomized in the slogan: “Science is the measure of all things, of those that are, that they are, and of those that are not, that they are not.” For modern mathematized natural science shorn of concern with laws, counterfactuals, and dispositions—in short of what is expressed by alethic modal vocabulary—is less than an impotent Samson; it is an inert, unrecognizable, fragmentary remnant of a once-vital enterprise. Sellars’s general recommendation for resolving this painful tension (felt particularly acutely by, and one of the principal issues dividing, the members of the Vienna circle) is to relax the exclusivism and rigorism he traces to empiricism’s semantic descriptivism (in a passage we have found reason to quote more than once already)<sup>3</sup>:

[O]nce the tautology ‘The world is described by descriptive concepts’ is freed from the idea that the business of all non-logical concepts is to describe, the way is clear to an *ungrudging* recognition that many expressions which empiricists have relegated to second-class citizenship in discourse are not *inferior*, just *different*.<sup>88</sup>

Sensitized as we now are by Sellars’s diagnoses of *semantic autonomy* claims as essential to various empiricist constructive and reconstructive projects, both in EPM and in the “Phenomenalism” essay, and familiar as we now are with his criticisms of them based on the inferentially articulated *doings* required to use or deploy various candidate base vocabularies, it should come as no surprise that his objections to critical empiricist suspicions of and hostility towards modality follow the same pattern. For the Humean-Quinean empiricist semantic challenge to the legitimacy of modal vocabulary is predicated on the idea of an independently and antecedently intelligible stratum of empirical discourse that is purely descriptive and involves no modal commitments, as a semantically autonomous background and model with which the credentials of modal discourse can then be invidiously compared.

---

<sup>88</sup> CDCM §79.



In this case, as in the others, the argument turns both on the *pragmatism* that looks to what one is doing in deploying the candidate base vocabulary—here “purely descriptive” vocabulary—and on the nature of the *inferential* articulation of that vocabulary necessary for such uses to play the expressive role characteristic of that vocabulary. The argument in this case is subtler and more complex than the others however. First, I take it that Sellars does *not* deny the intelligibility-in-principle of purely descriptive discourse that contains no explicitly modal vocabulary. Sellars is, frustratingly but characteristically, not explicit about his attitude towards the pragmatic autonomy in principle of such purely descriptive discourse. He says:

The idea that the world can, in principle, be so described that the description contains no modal expression is of a piece with the idea that the world can, in principle, be so described that the description contains no prescriptive expression. For what is being called to mind is the ideal of statement of ‘everything that is the case’ which, however, serves *through and through only* the purpose of stating what is the case. And it is a logical truth that such a description, however many modal expressions might properly be used in *arriving at* it or in *justifying* it, or in showing the *relevance* of one of its components to another, could contain no modal expression.<sup>89</sup>

Sellars’s view about this ideal is complex: there is sense in which it is intelligible, and a sense in which it is not. Such a discourse would be unreflective and unselfconscious in a way ours is not. For reasons that will emerge, it would belong to what at the end of the essay he calls the stage of human language “when linguistic changes had *causes*, but not *reasons*, [before] man acquired the ability to reason about reasons.”<sup>90</sup> (CDCM §108).

The second reason the argument must be subtler here is that there are special difficulties involved in, and corresponding delicacies required for, working out the general pragmatist-inferentialist strategy so as to apply it to this case, by specifying the relation between the

---

<sup>89</sup> CDCM § 80.

<sup>90</sup> CDCM §108.

expressive role distinctive of modal vocabulary, on the one hand, and what one is *doing* (in particular, the inferential commitments one is undertaking) in using ordinary, non-modal, descriptive vocabulary itself, on the other.

The pragmatic dependency relation that lies at the base of Sellars's argument is the fact that

although describing and explaining (predicting, retrodicting, understanding) are *distinguishable*, they are also, in an important sense, *inseparable*. It is only because the expressions in terms of which we describe objects, even such basic expressions as words for perceptible characteristics of molar objects, locate these objects in a space of implications, that they describe at all, rather than merely label. The descriptive and explanatory resources of language advance hand in hand.<sup>91</sup>

Descriptive uses of vocabulary presuppose an inferentially articulated “space of implications,” within which some descriptions show up as reasons for or explanations of others. Understanding those descriptions requires placing them in such a space. This pragmatist claim about what else one must be able to *do*—namely, *infer*, *explain*, treat one claim as a *reason* for another—in order for what one is doing to count as *describing* connects to the use of *modal* vocabulary via the principle that:

To make first hand use of these [modal] expressions is to be about the business of explaining a state of affairs, or justifying an assertion.<sup>92</sup>

That is, what one is *doing* in *using* modal expressions is explaining, justifying, or endorsing an inference. So what one is doing in saying that As are *necessarily* Bs is endorsing the inference from anything's being an A to its being a B.

---

<sup>91</sup> CDCM §108.

<sup>92</sup> CDCM §80.

The first sort of difficulty I alluded to above stems from the fact that there are other ways of endorsing such a pattern of inference besides *saying that* all As are necessarily Bs. One's endorsement may be *implicit* in other things one *does*, the reasoning one engages in and approves of, rather than *explicit* in what one *says*. So from the fact (assuming, as I shall, that it is a fact) that the activity of describing is part of an indissoluble pragmatic package that includes endorsing inferences and the fact that what one is doing in making a modal claim is endorsing an inference, it does not at all follow that there can be no use of descriptive vocabulary apart from the use of modal vocabulary. The second difficulty stems from the fact that although Sellars may be right that what one is *doing* in making a modal claim is endorsing a pattern of inference, it is clear that one is not thereby *saying that* an inference is good. When I say "Pure copper necessarily conducts electricity," and thereby unrestrictedly endorse inferences from anything's being pure copper to its conducting electricity, I have nevertheless *said* nothing about any inferences, explanations, justifications, or implications—indeed, have said something that could be true even if there had never been any inferences or inferencers to endorse them, hence no describers or discursive practitioners at all.<sup>93</sup> These two observations set the principal criteria of adequacy both for Sellars's positive working-out of the pragmatist-inferentialist treatment of modal vocabulary, and for his argument that the purely descriptive base vocabulary invoked by the empiricist critic of the semantic credentials of modal vocabulary lacks the sort of discursive autonomy the empiricist criticism presupposes and requires.

Sellars's central rhetorical strategy in this essay is to address the issue of what is expressed by modal claims about necessary connections by offering:

a sympathetic reconstruction of the controversy in the form of a debate between a Mr. C (for Constant Conjunction) and a Mr. E (for Entailment) who develop and

---

<sup>93</sup> Sellars connects this obvious fact with the observation that: Idealism is notorious for the fallacy of concluding that because there must be minds in the world in order for *us* to have reason to make statements about the world, therefore there is no sense to the idea of a world which does not include minds. (CDCM §101)

qualify their views in such a way as to bring them to the growing edge of the problem.<sup>94</sup>

Officially, he is even-handed in his treatment of the vices and virtues of the empiricist, who denies that the use of modal vocabulary can express any legitimate semantic content beyond that expressed by a descriptive, extensional universal generalization, and the rationalist, who understands that content in terms of entailments expressing rules of reasoning. In fact, however, as becomes clear when he launches into his own account, he is mainly concerned to develop a version of the rationalist account. As the second half of the essay develops, Sellars's marks his abandonment of the disinterested pose by an uncharacteristically explicit expository shift:

It is now high time that I dropped the persona of Mr. E, and set about replying to the challenge with which Mr. C ended his first critique of the entailment theory.<sup>95</sup>

Doing that requires careful investigation of the differences between and relations among four different sorts of item:

- Practical endorsement of the propriety of an inference from things being A to their being B;
- The explicit statement that one may infer the applicability of 'B' from the applicability of 'A';
- The statement that A physically entails B;
- The statement that As are necessarily Bs.

The first is the sort of thing Sellars takes to be pragmatically presupposed by the activity of describing, that is, deploying descriptive vocabulary. The second fails to capture such practical

---

<sup>94</sup> CDCM, introduction.

<sup>95</sup> CDCM §85. In fact, Sellars's 'defense' of Mr. C (see the passage from § 82 quoted in note 3 above) consists of showing what concessions he needs to make to Mr. E. This proceeds first by Mr. C's qualification that "'A causes B' says that  $(x)[Ax \rightarrow Bx]$  and *implies* that the latter is asserted on inductive grounds" (CDCM §62), followed by the necessity of conceiving "of induction as establishing principles *in accordance with which* we reason, rather than as major premises *from which* we reason." (CDCM §83) As will appear, the former concession, introducing the notion of what is contextual implied by contrast to what is explicitly said, is then dialectically made available to be pressed into service by Mr. E. This bit of dialectic is a pretty rhetorical flourish on Sellars's part, but I doubt that in the end it reflects any deep feature of the confrontation between the empiricist and rationalist approaches to modality.

endorsements, because of the possibility of asserting such statements regarding the *expressions* ‘A’ and ‘B’ without understanding what they express.<sup>96</sup>

The third sort of statement expresses Mr. E’s initial stab at an analysis of the fourth. It is the answer to the question: what sort of entailment is it that modal statements are supposed to express?

Mr. E has a ready answer. ... it might ... be called ‘natural’ or ‘physical’ entailment, for while any entailment is a logical relation, we can distinguish within the broad class of entailments between those which are, and those which are not, a function of the specific empirical contents between which they obtain. The latter are investigated by general or formal logic (and pure mathematics). Empirical science, on the other hand, to the extent that it is a search for *laws*, is the search for entailments of the former kind. (Putative) success in this search finds its expression in statements of the form ‘It is (inductively) probable that A physically entails B.’<sup>97</sup>

The virtue of statements like “A physically entails B” is that they do plausibly codify the practical endorsement of an inference that is implicit in what one does in the form of something one can explicitly *say*, without bringing in irrelevant commitments concerning particular expressions, the activity of inferring, or discursive practitioners. The remaining difficulty is that they seem plainly not to have the same content, not to say the same thing, as explicitly modal statements of objective necessity.

Sellars’s response to this problem is to acknowledge that modal statements do not *say that* some entailment holds, but to distinguish between what is *said* by using a bit of vocabulary

---

<sup>96</sup> As Sellars says:

But one can know that Turks, for example, ought to withdraw ‘...’ when they commit themselves to ‘---’ without knowing the language, whereas the statement that ‘*p* entails *q*’ contextually implies that the speaker not only knows the language to which ‘*p*’ and ‘*q*’ belong, but, in particular, knows how to use ‘*p*’ and ‘*q*’ themselves. (CDCM § 81)

<sup>97</sup> CDCM §56.

and what is ‘*contextually implied*’ by doing so. Sellars says very little about this latter notion, even though it bears the full weight of his proposed emendation of the rationalist account. It is recognizably the same distinction he had appealed to earlier, in “Inference and Meaning”, as the distinction between what one *says* by making a statement and what one thereby *conveys*. There his example is that in asserting “The sky is clear today,” I *say* that the sky is clear today, but *convey* that I *believe* that it is clear.<sup>98</sup> That otherwise uninterpreted example suggests to me that what Sellars has in mind is the distinction between *semantic* and *pragmatic* inferences. That is the distinction between inferences underwritten by the *contents* of what is *said* or asserted, on the one hand, and inferences underwritten by what one is *doing* in saying them, on the other. The inference from “The sky is clear” to “It is not raining” is of the first sort; the inference from my asserting “The sky is clear” to “Brandom believes the sky is clear” is of the second sort. Inferences of these two kinds may generally be distinguished by the Frege-Geach embedding test: look to see whether those who make the inference in question also endorse the corresponding conditional. “If the sky is clear, then it is not raining” is generally true, while “If the sky is clear, then Brandom believes it is clear” is not generally true. (Compare the inference from my *saying* “That is an ugly tie you are wearing” to “Bob is annoyed with me.”)

---

<sup>98</sup> Sellars, “Inference and Meaning”, p. 280/332 in J. Sicha (ed.) *Pure Pragmatics and Possible Worlds: The Early Essays of Wilfrid Sellars* [Ridgeview Publishing Company, Reseda CA, 1980].

## V. Kantian Pragmatism about Modality

If that is in fact the distinction Sellars is after, then it seems to me that the view he is expounding and defending can be put less paradoxically if we do not take a detour through entailment statements, but concern ourselves directly with the relation between the endorsement of patterns of inference and modal statements. The underlying rationalist insight is a pragmatist-inferentialist one: what one is *doing* in making a modal claim is endorsing a pattern of inference. Modal vocabulary makes possible new kinds of *sayings* that have the *pragmatic effect* of endorsing inferences. To say that is not yet to say what they *say*, it is only to say what one is *doing by saying* them. But it does settle the *pragmatic significance* of such modal claims, in the sense of their appropriate circumstances and consequences of application.<sup>99</sup> If one practically endorses the pattern of inference that treats classifying or describing anything at all as an A as sufficient grounds (“all on its own”, as Sellars says, in order to capture the way the pattern of inferences in question is counterfactually robust) for concluding that it is a B, then one is committed to the claim that all As are necessarily Bs. And commitment to that claim is

---

<sup>99</sup> It is the attempt to specify this peculiar and distinctive sort of pragmatically mediated relation between vocabularies that leads Sellars to say things like:

It is sometimes thought that modal statements do not describe states of affairs in the world, because they are *really* metalinguistic. This won't do at all if it is meant that instead of describing states of affairs in the world, they describe linguistic habits. It is more plausible if it is meant that statements involving modal terms have the force of *prescriptive* statements about the use of certain expressions in the object language. Yet there is more than one way of to '*have the force of*' a statement, and failure to distinguish between them may snowball into a serious confusion as wider implications are drawn. (CDCM § 81)

and

Shall we say that modal expressions are metalinguistic? Neither a simple 'yes' nor a simple 'no' will do. As a matter of fact, once the above considerations are given their proper weight, it is possible to acknowledge that the idea that they are metalinguistic in character oversimplifies a fundamental insight. For our present purposes, it is sufficient to say that the claim that modal expressions are 'in the metalanguage' is not too misleading if the peculiar force of the expressions which occur alongside them (represented by the 'p' and the 'q' of our example) is recognized, in particular, that they have 'straightforward' translation into other languages, and if it is also recognized that they belong not only 'in the metalanguage', but in discourse about *thoughts* and *concepts* as well. (CDCM § 82)

and

We must here, as elsewhere, draw a distinction between what we are committed to concerning the world by virtue of the fact that we have reason to make a certain assertion, and the force, in a narrower sense, of the assertion itself. (CDCM §101)

commitment to practically ratify that pattern of inference. Assuming, as Sellars has claimed, that using ordinary, non-modal, descriptive vocabulary requires practically endorsing such patterns of inference (“situating descriptions in a space of implications”), anyone who has the practical ability to deploy “purely descriptive” vocabulary already knows how to do everything he needs to know how to do to deploy modal vocabulary as well. He need not actually do so, since practically undertaking those inferential commitments does not require that one have available a language with vocabulary permitting one to *do* that by *saying* something. But *all* a practitioner lacks in such a circumstance is the *words* to hook up to discriminative and responsive abilities he already possesses. In this precise sense, the ability to deploy modal vocabulary is *practically implicit* in the ability to deploy non-modal descriptive vocabulary.

Sellars has claimed that the activity of describing is unintelligible except as part of a pragmatic package that includes also not just the making of inferences, but the making of *counterfactually robust* inferences: the sort of inferences involved in *explanation*, and licensed by explicitly modal statements of *laws*. He sums up the claim admirably in the title of another one of his earliest papers: “Concepts as Involving Laws, and Inconceivable without Them.”<sup>100</sup> Grasp of a concept is mastery of the use of a word, Sellars says. And for descriptive concepts, that use includes not only sorting inferences (however fallibly and incompletely) into materially good and materially bad ones, but also, among the ones one takes to be materially good, to distinguish (however fallibly and incompletely) between counterfactual circumstances under which they do, and counterfactual circumstances under which they do not, *remain* good. Part of taking an inference to be materially good is having a view about which possible additional collateral premises or auxiliary hypotheses would, and which would not, infirm it. Chestnut trees produce chestnuts—unless they are immature, or blighted. Dry, well-made matches strike—unless there is no oxygen. The hungry lioness would still chase the antelope if it were Tuesday or the beetle on the distant tree crawled slightly further up the branch, but not if lioness’s heart were to stop beating. The point is not that there is any particular set of such discriminations that one must be able to make in order to count as deploying the concepts involved. It is that if one can make *no* such practical assessments of the counterfactual

---

<sup>100</sup> In *Pure Pragmatics and Possible Worlds*, ed. Jeffrey Sicha (Atascadero, CA: Ridgeview, 1980), pp. 87-124.



robustness of material inferences involving those concepts, one could not count as having mastered them.

Against the background of this pragmatist-inferentialist claim about what is involved in the ordinary descriptive use of concepts, Sellars's claim, as I am reading him, is that explicitly modal "lawlike" statements are statements that one is committed or entitled to whenever one is committed or entitled to endorse such patterns of counterfactually robust inference, and commitment or entitlement to which in their turn commit or entitle one to the corresponding patterns of inference. Saying that about them settles what one needs to *do to use* such modal statements. It does *not* say how one is thereby *describing* the world as being when one does. It does not, in particular, *describe* a pattern of inference as good (though that saying does, in its own distinctive way, *express endorsement* of such a pattern). It does not do those things for the simple reason that the use of modal expressions is *not* in the first instance *descriptive*.<sup>101</sup> It codifies explicitly, in the form of a statement, a feature of the use of descriptive expressions that is indissolubly bound up with, but not identical to, their descriptive use. Nonetheless, in knowing how to use vocabulary descriptively, one knows how to do everything one needs to know how to do in order to use modal vocabulary. And that is enough to show that one cannot actually be in the Humean predicament presupposed by the empiricist challenge to the intelligibility of modal vocabulary. For one cannot know how to use vocabulary in matter-of-factual descriptions ("The cat is on the mat") and not have any grip on how to use modal, counterfactual, and dispositional vocabulary ("It is necessary for live cats to breathe," "The cat could still be on the mat if the mat were a slightly different shade of blue, but not if it turned into soup," "The cat would leave the mat if she saw a mouse."). Although *explicitly* modal *vocabulary* is an in-principle optional superstructure on practices of deploying descriptive vocabulary, what it expresses cannot be mysterious in principle to those who can engage in those base-level practices.

---

<sup>101</sup> Sellars says:

[Mr. E.] conceives of induction as establishing principles *in accordance with which* we reason, rather than as major premises *from which* we reason. [CDCM §83]

In taking this line, Sellars quite properly sees himself as reviving a central idea of Kant's. The ability to use empirical descriptive terms such as 'mass', 'rigid', and 'green' already presupposes grasp of the kind of properties and relations made explicit by modal vocabulary. It is this insight that leads Kant to the idea of 'pure' concepts or 'categories', including the alethic modal concepts of necessity and possibility that articulate causal laws, which must be available *a priori* because and in the sense that the ability to deploy them is presupposed by the ability to deploy ordinary empirical descriptive concepts. The categories, including modality, are concepts that make explicit what is implicit in the empirical, descriptive use of any concepts at all. Though the details of *which* laws, the statements of which express counterfactually robust patterns of inference, actually obtain is an empirical one, *that* empirical descriptions are related by *rules* in the form of laws, which do support counterfactually robust inferences, is *not* itself an empirical matter, but a truth about the framework of empirical description. I want to call the underlying insight "the Kant-Sellars thesis about modality." It is the claim that in being able to use non-modal, empirical-descriptive vocabulary, one already knows how to do everything one needs to know how to do in order to deploy modal vocabulary, which accordingly can be understood as making explicit structural features that are always already implicit in what one *does* in describing.

## VI. Conclusion

Articulating and justifying his version of the Kant-Sellars thesis about modality is Sellars's constructive response to the empiricist tradition's "nothing-but-ism" about modality: its demand that what is expressed by modal claims either be shown to be expressible in non-modal terms, or be dispensed with entirely by semantically fastidious philosophers and scientists. This complements and completes his demonstration, in the "Phenomenalism" essay, that this critical consequence of an over-ambitious empiricism is in any case incompatible with any constructive empiricist effort to reconstruct or replace the use of target vocabularies such as objective-descriptive vocabulary, primary-quality vocabulary, and theoretical vocabulary in terms of the favored empiricist base vocabularies, if that effort is subject to even the most minimal criteria of material adequacy. Together, these arguments show what Sellars eventually made of his early intuition that the soft underbelly of empiricism, in both its traditional and its twentieth-century logistical form, is its semantic treatment of modality.

My overall aim in this lecture has been to place the arguments against empiricism presented in the first half of "Empiricism and the Philosophy of Mind" in the larger context opened up by laying them alongside the further battery of arguments aimed at the same target that derive from consideration of that tradition's views about modality. And I have been concerned to show that the methodological strategies that guide all of these discussions are Sellars's *pragmatist* insistence on looking at what one must be able to *do* in order to deploy empirical descriptive vocabulary, and his *rationalist* commitment to the necessary *inferential* articulation of the concepts expressed by the use of such vocabulary. I think that even fifty years on, there is still a lot of juice to be squeezed out of these ideas.

But I want to close with another, perhaps more frivolous suggestion. Every sufficiently engaged reading becomes a rewriting, and I have been offering here, *inter alia*, the outline of a

different narrative strategy that Sellars could have adopted in the late 1950s. Under some such title as *The Limits of Empiricism*, he could have re-presented the material that in fact appeared first as roughly the first half of “Empiricism and the Philosophy of Mind,” and the second halves of each of “Phenomenalism” and “Counterfactuals, Dispositions, and Causal Modalities,” organized around and introduced in terms of the themes I have traced here. It is interesting to speculate about how his reception might have been different—and about where we would find ourselves today—had this been the shape of Sellars’s first book.

## 2014 Nordic Pragmatism Lectures:

### Analytic Pragmatism, Expressivism, and Modality

#### Lecture 5

#### Modal Expressivism and Modal Realism: Together Again

##### I. A Modal Expressivism

Kant saw that in addition to concepts whose principal use is to make it possible for us to describe how things are, there are concepts that make explicit features of the metaconceptual *framework* that makes such description possible. An important class of the framework-explicating concepts (arguably the one that motivated this entire line of thought) comprises *alethic modal* concepts, such as necessity and possibility. These express lawful relations between ground-level descriptive concepts, and mark the special status of Newton's laws, their lawfulness, by contrast to the status of merely contingent matters of fact, the role played by statements of initial and boundary conditions for particular applications of those laws. But it is not only in understanding the use of technical scientific concepts that the modal concepts find application. The use of ordinary empirical descriptive concepts such as gold, and cat, and house, no less than the Newtonian concepts of mass, force, and acceleration, is essentially, and not just accidentally, articulated by the modality these modal concepts express.

It is because he believes all this that Kant calls modal concepts (among others) 'pure' concepts: categories. Pure concepts are a species of *a priori* concepts.<sup>102</sup> The sense in which we

---

<sup>102</sup> That is, concepts available *a priori*. I take it that Kant's standard usage of "a priori" is adverbial, though this is not obvious since the Latin phrase is not grammatically marked as it would be in German. *Exactly* what Kant means by the term 'pure' [rein], as it applies generically to reason, knowledge, understanding, principles, concepts, and intuition is a complex and challenging question. There seems to be some terminological drift across the species, and some wavering on how to classify particular examples. (The status of the crucial *a priori* principle that every

can think of them as available *a priori* that I want to focus on comprises three claims. First, what they express are structural features of the framework within which alone it is possible to apply *any* concepts, make *any* judgments, including ordinary empirical descriptive ones. Second, in being able to apply any ground-level empirical concepts, one already knows how to do everything one needs to know how to do in order to apply the categorial concepts. Finally, there are no *particular* empirical descriptive concepts one must be able to apply in order to have implicit mastery of what is expressed by categorial concepts such as the modal ones (though perhaps one must have some descriptive concepts or other).

The alethic modality that has this categorial status is something like physical necessitation. It is the modality involved in the “pure principle” that “every alteration must have a cause.” But the use of these modal concepts to formulate particular laws of nature results neither in *a priori* principles nor in analytic judgments. Lawlike claims assert modal relations between noncategorial descriptive concepts. They are synthetic, and must be discovered and justified empirically. The crux of Kant’s challenge in the first Critique that culminates in the B Deduction, is to show how it is intelligible that categorial concepts, paradigmatically the modal ones, can both articulate structural relations intrinsic and essential to the use of descriptive *concepts* and express *causal* laws of nature that combine the features of being on the one hand universal and necessary, and on the other hand, empirical.

A further development of what I want to claim will be retrospectively recognizable as the same line of thought can be found in Frege.<sup>103</sup> His use of Latin letters and his logical sign of generality (used in conjunction with the notation for hypotheticals) express relations between concepts. It has always been an embarrassment for the anachronistic extensional quantificational reading of this notation (due originally to Russell) that Frege says of it, when he first introduces

---

alteration must have a cause, for instance, is apparently variously characterized at [B3] and [B5].) Being available *a priori* is necessary, but not sufficient [B3].

103 The characterization of Frege’s Begriffsschrift that follows is one that I had my eyes opened to by Danielle Macbeth’s pathbreaking book *Frege’s Logic* [Harvard University Press, 2005].

it in the *Begriffsschrift*, that it is the right way to express *causal* relations of necessitation.<sup>104</sup> For it is a commonplace of the later logistical tradition that merely quantificational relations between concepts cannot distinguish between contingent regularities and lawlike, necessary ones. For that, explicit modal operators must be applied to the quantified conditionals.

But Frege deploys his notation so that the relations between concepts expressed by generalized conditionals *already* have modal force. Relations between concepts of the sort logic lets us express have consequences for relations between their extensions, of the sort our quantificational notation expresses, but his generality locutions (the use of Latin letters and the concavity with German ones) codify relations we think of as intensional. Fregean logical concepts are indeed second- and higher-order concepts, but more than that, the universality they express is *rulish*. They are in the first instance principles in accordance with which to reason, and only derivatively premises from which to reason.<sup>105</sup> In addition to permitting the formulation of purely logical relations among logical concepts, Frege's logical vocabulary permits us to assert necessary connections among empirical concepts that themselves can only be discovered empirically: physically or causally necessary connections. In the Preface to the *Begriffsschrift*, Frege says:

It seems to me to be easier still to extend the domain of this concept-script [Begriffsschrift] to include geometry. We would only have to add a few signs for the intuitive relations that occur there...The transition to the pure theory of motion and then to mechanics and physics could follow at this point. The latter two fields, in which besides rational necessity [Denknotwendigkeit] natural necessity

---

104 “*This is the way in which causal connections are expressed.*” [Italics in the original.] *Begriffsschrift* §12 (p. 27 in Jean van Heijenoort (ed.) *From Frege to Gödel: A Source Book in Mathematical Logic, 1879-1931* [Harvard University Press, 1967]), foreshadowed at §5.

105 Following Mill, this is Sellars's way of putting the point, in “Counterfactuals, Dispositions, and the Causal Modalities” Pp. 225-308 of *Minnesota Studies in the Philosophy of Science*, Vol. II, ed. by H. Feigl, M. Scriven, and G. Maxwell, (University of Minnesota Press; Minneapolis, MN: 1957). Henceforth “CDCM.”

[Naturnotwendigkeit] asserts itself, are the first for which we can predict a further development of the notation as knowledge progresses.<sup>106</sup>

The additional signs that such an extension requires do *not* include modal operators. The necessity (whether natural or rational) of the connections between empirical concepts is already contained as part of what is expressed by the *logical* vocabulary, even when it is used to make claims that are not logically, but only empirically true.

The capacity to express modal connections of necessitation between concepts is essential to Frege's overall purpose in constructing his *Begriffsschrift*. Its aim is to make explicit the contents of concepts. Frege understands that content as articulated by the *inferential* relations between concepts, and so crafted his notation to make those inferential connections explicit. Introducing his project in the third section of the *Begriffsschrift*, he says:

The contents of two judgments may differ in two ways: either the consequences derivable from the first, when it is combined with certain other judgments, always follow also from the second, when it is combined with the same judgments, or this is not the case. The two propositions "The Greeks defeated the Persians at Plataea," and "The Persians were defeated by the Greeks at Plataea," differ in the first way...I call that part of the content that is the *same* in both the *conceptual content* [begrifflich Inhalt]...[I]t alone is of significance for my concept-script [Begriffsschrift].

The principal technical innovation that makes it possible for the *Begriffsschrift* to express the inferential relations that articulate conceptual content, Frege takes it, is his notation for generality, when used in connection with his conditional (used to express hypothetical judgeable contents). An essential element of that expressive power is the capacity of this notation to express *robustly*, *modally robust*, inferential relations of necessitation, including, importantly, the natural necessity characteristic of inferences underwritten by *causal* connections. Though he doesn't himself think of it this way, Frege is continuing and developing Kant's line of thought concerning the role that modality (including centrally the kind of necessity involved in causation) plays in distinguishing the expressive role of certain concepts that relate ground-level empirical descriptive concepts to one another from the expressive role of those descriptive concepts themselves.

Nearer to our own time, this line of thought has been further developed and clarified by Wilfrid Sellars. He lucidly compressed his endorsement of the fundamental Kantian idea that modal concepts make explicit something implicit in the use of ordinary empirical descriptive

---

106 P. 7 in van Heijenoort op.cit.. I have emended the translation slightly, where I have noted the original German terms.



concepts into the title of one of his earliest essays: “Concepts as Involving Laws, and Inconceivable without Them.” But he also offers the outline of a more articulated argument for the claim. We can reconstruct it as follows:

1. “It is only because the expressions in terms of which we describe objects... locate these objects in a space of implications, that they describe at all, rather than merely label.”<sup>107</sup>
2. It is an essential feature of the inferential relations in which, according to claim (1), descriptive concepts must stand, that they can be appealed to in *explanations* and *justifications* of further descriptions.
3. So: “although describing and explaining (predicting, retrodicting, understanding) are *distinguishable*, they are also, in an important sense, *inseparable*... The descriptive and explanatory resources of language advance hand in hand...”<sup>108</sup>
4. The expressive role distinctive of modal vocabulary is to make explicit these explanatory and justificatory relations.

This line of thought is a way of filling in ideas that Sellars had had since his student days. In an autobiographical sketch, he tells us that he was to begin with concerned to understand the sort of content expressed by concepts of the “logical, causal, and deontological modalities.” (Here only what he calls the “causal” modalities are at issue—a point to which I shall return.) His big idea, he tells us, was that

what was needed was a functional theory of concepts which would make their role in reasoning, rather than supposed origin in experience, their primary feature.<sup>109</sup>

The idea he got from Kant was that the “role in reasoning” distinctive of a key class of alethic modal concepts is to articulate the “role in reasoning” of ordinary empirical descriptive concepts.

---

107 “Counterfactuals, Dispositions, and Causal Modalities” In *Minnesota Studies in the Philosophy of Science, Volume II: Concepts, Theories, and the Mind-Body Problem*, ed. Herbert Feigl, Michael Scriven, and Grover Maxwell (Minneapolis: University of Minnesota Press, 1958), p.225-308.] (hereafter CDCM), § 108.

108 CDCM § 108.

109 In *Action, Knowledge, and Reality*, H. N. Castaneda (ed.) [Indianapolis, Bobbs-Merrill, 1975] p 285.

The two key moves in an argument of this form are, first, an account of the descriptive use of empirical concepts that exhibits as essential their articulation by inferences that can support explanations and justifications, and second, an account of the central function of at least some alethic modal vocabulary as expressing explanatory and justificatory inferential relations among descriptive concepts. The conclusion of the argument is what I call the “Kant-Sellars thesis about modality”: in knowing how to use ordinary empirical descriptive vocabulary, one already knows how to do everything one needs to know how to do in order to be able (in principle) to use alethic modal vocabulary.<sup>110</sup> According to this thesis, one cannot be in the semantic predicament that empiricists such as Hume and Quine envisaged: understanding ordinary empirical descriptive vocabulary perfectly well, but having thereby no grip at all on what is expressed by modal vocabulary.

How does Sellars understand the distinction between “merely labeling”, on the one hand, and describing, in the sense he then wants to argue “advances hand in hand” with explaining and justifying, on the other hand? Labeling is attaching signs to, or associating them with, items in the nonlinguistic world. The paradigm of this semantic relation is that between an arbitrary name and its bearer, or a sign and what it signifies—what Sellars elsewhere calls “the ‘Fido’-Fido model.” Now it is one of the founding insights of analytic philosophy of language that the results of a Procrustean assimilation *all* semantic relations to this nominalistic model are disastrous. That is a lesson taught originally by Frege, and again by both the Wittgenstein of the *Tractatus* and the Wittgenstein of the *Investigations*, each in his own way. (The mistake lives on in semiotics and in the structuralist heirs of de Saussure. Derrida was sufficiently in the grip of this traditional picture that the only alternative to it he could conceive was that signs should be understood to stand exclusively for...other signs.) What one will not understand on this model, in the first instance, is what is special about *sentences*, and what they express: claimables, judgeable contents, Fregean thoughts as thinkables. In particular, using the ‘Fido’-Fido model to think about the relation between declarative sentences and *true* Fregean thinkables, facts, is fraught with difficulties. Indeed, even the more promising strategy that avoids the nominalistic mistake of modeling the semantics of sentences on that of names while crafting a technical notion of *representation* to be generic across its disparate name-bearer and (true) sentence-fact species requires more subtlety, craft, and guile than is generally appreciated.

---

110 I discuss this claim at greater length in Chapter Four of *Between Saying and Doing: Towards an Analytic Pragmatism* (Oxford: Oxford University Press, 2008).

Of course, one need not make the nominalistic mistake of assimilating *all* semantic relations to labeling in order to claim that the model applies to *some* uses of linguistic expressions, that is, to claim that there are, after all, labels—even if sentences are not to be counted among them. Sellars is claiming that describing should also not be assimilated to applying a “mere label.” Here the relevant grammatical category is not terms or sentences, but predicates. Predicate labels in Sellars’s sense can have more content than proper names like ‘Fido’. The use of predicates to make observation reports requires the user to exercise a reliable differential responsive disposition. It is tempting to think that reliably responding in a distinctive way to some things and not others is a way of *classifying* them as being of some kind, or as having something in common. What more besides dividing things into groups could be required to count as *describing* them as being of different kinds? The difference between classifying in the sense of labeling and describing emerges when we ask what the things grouped together by their elicitation of a common response are supposed to be described *as*. If the dog reliably barks at some things, and not others (cats, dogs, and squirrels, but not horses, men but not women, motorcycles but not cars, helicopters but not airplanes, church bells but not the neighbor’s stereo, and so on) it is *grouping* things, sorting them into two classes. But there need be nothing it is *describing* them *as*. When the metal strip expands in some environments and contracts in others, it is not yet *describing* them as warm or cold.

Sellars’s idea is that what one is describing something *as* is a matter of what *follows* from the classification—what *consequences* falling in one group or another has. It is insofar as being grouped one way rather than another can serve as a *premise* in an *inference* that the grouping is intelligible as a *description* and not merely a label. Even in the primitive, noninferential case of the three vervet cries appropriately elicited (as the young ones are trained by their elders) by snakes, eagles, and leopards, it is insofar as they are appropriately responded to (as the young ones are trained by their elders) by jumping, covering, and climbing respectively that they begin to be intelligible as describing threats-from-below, threats-from-above, and so on. Reliably differentially elicited responses are intelligible as observation reports, as empirical descriptions, just insofar as they are available to *justify* further claims. It is essential, and not just accidental, to descriptive predicates that they can be used to make claims, which would be expressed by declarative sentences. And it is essential, and not accidental to those claimings that they can serve as *reasons* for further claims. (Of course, this Sellarsian inferentialist way of developing Frege’s claims about how we must think of the contents of predicates and sentences as related to one

another once we see the inadequacy of nominalistic construals is controversial. I have elaborated and defended it elsewhere, and am merely expounding it here.)

In the same spirit, Michael Dummett argues that the content of a descriptive concept cannot be identified with its circumstances of appropriate application alone. In order to avoid the defects and inadequacies of one-sided theories of meaning, one must consider *both* those circumstances of application *and* the appropriate consequences of such application—which is to say also its role as a premise in inferences (both theoretical and practical). It is possible to construct descriptive concepts that share circumstances or consequences of application, but differ in the other component. In such cases, they differ also in their content or meaning. Thinking of the application of substantive nonlogical descriptive concepts as involving a commitment to the propriety of the material inference from their circumstances to their consequences of application is a way of insisting that descriptive concepts count as locating the objects they are applied to “in a space of implications.”

Sellars sees modal locutions as tools used in the enterprise of

...making explicit the rules we have adopted for thought and action...I shall be interpreting our judgments to the effect that A causally necessitates B as the expression of a rule governing our use of the terms ‘A’ and ‘B’.<sup>111</sup>

The rules they express are rules of *inference*. Modal expressions are inference licenses or inference “tickets,” in Ryle’s terminology.<sup>112</sup> These are what Sellars calls “material,” that is, non-logical inferences. In fact, what these modal locutions make explicit, according to Sellars, are just the implications, situation in a space of which is what distinguishes descriptive concepts from mere labels. Inferences such as that from “Pittsburgh is to the West of Princeton, so Princeton is to the East of Pittsburgh,” articulate the content of the descriptive concepts West and East.

---

111 Sellars, "Language, Rules, and Behavior" footnote 2 to p. 136/296 in *PPPW*.

112 Gilbert Ryle, “‘If’, ‘So’, and ‘Because’ ”, pp. 302-318 in Black, Max (ed.) *Philosophical Analysis* [Prentice Hall, 1950]. Sellars does not discuss whether “A causally necessitates B” should be understood as expressing a committive, or merely a permissive inference.

Further, it is the inferential commitments acknowledging such material implicational relations that are appealed to in explanation and justification.

To make first hand use of these [modal] expressions is to be about the business of explaining a state of affairs, or justifying an assertion.<sup>113</sup>

That is, what one is *doing* in *using* modal expressions (“As are necessarily Bs”) is endorsing an inference (from anything’s being A to its being B) that can be appealed to in justifying one description on the basis of another, or explaining the applicability of one description by the appealing to the applicability of another: “The raspberries are red *because* they are ripe.” This is why the expressive resources of description, on the one hand, and justification and explanation, on the other hand, “advance hand in hand,” as Sellars says.

Because he understands the expressive function characteristic of the modal vocabulary he is addressing to be that of making explicit the inferential relations appealed to in justifications and explanations, Sellars takes it that the central use of that vocabulary is in qualifying conditionals, paradigmatically quantified conditionals, rather than their use as operators applying to nonconditional descriptive sentences. What the modal vocabulary expresses is the element of *generality* that Ryle had insisted was present in all endorsements of inferences:

...some kind of openness, variableness, or satisfiability characterizes all hypothetical statements alike, whether they are recognized “variable hypotheticals” like “For all  $x$ , if  $x$  is a man,  $x$  is mortal” or are highly determinate hypotheticals like “If today is Monday, tomorrow is Tuesday.”<sup>114</sup>

That element of generality would naturally be made explicit in this last example by applying a necessity operator to the conditional. Another way of putting this same point is that the inferential relations among descriptive concepts in virtue of which they can be used to *describe*, and not just *label*, which are appealed to in *justifications* and *explanations* of the applicability of one description on the basis of the applicability of another, and which are made explicit by the use of modally qualified conditionals are *subjunctive* and *counterfactual* supporting inferences. They make explicit the laws that Sellars says concepts involve and are inconceivable without.

---

113 CDCM § 80.

114 Gilbert Ryle “‘If’, ‘So’, and ‘Because’”, pp. 302-318 in Black, Max (ed.) *Philosophical Analysis* [Prentice Hall, 1950], p. 311.

This constellation of claims to which Sellars aspires to entitle himself articulates what he makes of the tradition of thinking about modality that Kant initiates and Frege develops in an inferentialist key. It is a story that construes (at least one kind of) modal vocabulary as distinguished by the role it plays in expressing explicitly essential aspects that it makes visible as implicit already in the use of ordinary empirical descriptive vocabulary. Having a (“first hand”) use in explicating the framework within which vocabulary use can have the significance of describing—a framework we come to see as necessarily a unified package comprising not only description, but justification and explanation, a framework articulated by subjunctively robust inferential relations among descriptive concepts—sets modal vocabulary off from the descriptive vocabulary, precisely in virtue of the distinctive expressive role it plays with respect to the use of such descriptive vocabulary. This, then, is Sellars’s modal expressivism.

It is, it should be acknowledged, largely programmatic. Turning the program into a full-blooded account of the use of modal vocabulary would require satisfactory responses to a number of challenges. I remarked above that Sellars’s approach focuses on modally qualified conditionals. So, at a minimum, we would need to understand how it might be developed or extended to deal with other uses of modal operators.<sup>115</sup>

A second issue concerns the kind of modality Sellars is telling us about. His topic patently is not *logical* necessity and possibility. Nor is it the sort of *metaphysical* necessity and possibility Kripke introduces us to in “Naming and Necessity.” In the principal essay in which he develops his expressivism, Sellars specifies what he is interested in as “causal” modalities.<sup>116</sup> There and elsewhere he talks about them as “physical” modalities. It is clear that he means to be discussing the sort of alethic necessity and possibility that characterizes laws of nature—not only laws of fundamental physics, but also laws promulgated in the special sciences. He seems to think that this is generically the same modality as that involved in ordinary informal explanations of empirical phenomena: of why the car wouldn’t start, why the beans burned, why the squirrel couldn’t get to the bird-feeder, and so on. It is clearly some such notion of necessity and possibility that Kant was addressing. It is the kind of necessity that is the target of Hume’s skeptical epistemological doubts about the possibility of establishing on inductive grounds, and of his consequent semantic doubts about its ultimate intelligibility. Frege’s few, gnomic remarks about the modal force of

---

115 Semantic inferentialists think that the use of *any* concept involves commitment to the propriety of all the inferences from the circumstances of appropriate application to the appropriate consequences of application of that concept. (Cf. Chapter One of *Articulating Reasons* [Cambridge: Harvard University Press, 1997]. So in that context, a strategy for addressing this challenge might not be far to seek.

116 CDCM.

his generality locutions (the concavity and the use of latin letters) suggest he was thinking about something like this same notion of necessity.

Sellars also clearly thinks that it is a kind of *conceptual* necessity. The modality he is analyzing characterizes the subjunctively robust inferential connections among empirical concepts in virtue of which (at least in part) they have the descriptive contents that they do. The laws, exhibiting that modality, which such concepts involve (without which, we are told, they are inconceivable) articulate the contents of those concepts, or at least the framework within which they are intelligible as having those contents. This aspect of Sellars's thought is what he makes of Kant's treatment of alethic modality as a *category*, a *pure* concept. For those, Sellars thinks, are the concepts that make explicit what something implicit in the use of any empirical descriptive concepts. This is the *semantic* sense in which they are always available *a priori*: apart from the applicability of any *particular* noncategorical, empirical concepts.

But it is not easy to see how to reconcile these two characterizations of the modality in question: as causal, physical necessity and possibility, and as some sort of conceptual necessity and possibility. In particular, these two conceptions of a kind of alethic modality seem to pull in different directions epistemologically. For laws of nature, or statements about what causally or physically necessitates what (or makes what else causally or suggestion possible or impossible) must in general be established empirically. But questions of what is conceptually necessary or possible, of what other concepts must or can be applied if some concept *were* to be applied, just in virtue of the contents of the concepts involved, seems to be something one can discover *a priori*. One does not need to know how the world is, only what one means—not what descriptive concepts actually apply to a situation, but only what the contents of those concepts are. We are faced with an inconsistent triad of a form that is familiar to readers of “Empiricism and the Philosophy of Mind”<sup>117</sup>:

- Physical or causal necessity and possibility are a kind of conceptual necessity and possibility.
- Physical or causal necessities and possibilities must be established empirically.
- Conceptual necessities and possibilities can be established *a priori*.

---

117 Edited by Robert Brandom, with an Introduction by Richard Rorty [Cambridge: Harvard University Press [ref.]] §[6 or 7, check]. Notice that insofar as there is any go to Sellars's reading of Kant on this point, a corresponding issue arises for Kant's view. How is it, exactly, that we can know *a priori* that nature is lawful, but can only know empirically what the laws are? I say something about this issue in Section IV.

Sellars is fully aware of this difficulty, and has a straightforward, if radical, response. He rejects the third element of the triad. A semantic externalist *avant la letter*, he takes it that we cannot discover the contents of our concepts or the meanings of our words just by introspecting. He follows Kant in understanding concepts as rules (norms) we bind ourselves by, without knowing everything about what we are committing ourselves to by applying those concepts. Finding out what applications of descriptive concepts are correct and finding out what inferences connecting those descriptive concepts are correct are two sides of one coin, two aspects of one process of empirical inquiry. Though Quine would not put the point this way, Sellars is at one with him in denying the Carnapian two-phase story (appropriate for formal languages, but not for natural languages) according to which first, by one sort of procedure one has privileged, nonempirical access to, one fixes meanings (concepts, the language) and then subsequently, by another sort of procedure, which is empirical, determines the facts (what to believe, one's theory) as expressed in those meanings (concepts, language). To find out what the contents of the concepts we apply in describing the world really are, we have to find out what the laws of nature are. And that is an empirical matter.

Another challenge to working out Sellars's version of modal expressivism concerns the extent to which, and the sense in which, it should be understood as taking the expressive role characteristic of modal vocabulary to be a *metalinguistic* one. On the one hand, when Sellars says he wants to understand a paradigmatic kind of modal judgment as "the expression of a rule governing our use of the terms 'A' and 'B'," this sounds straightforwardly metalinguistic in a classical sense. (This formulation is from an early paper, and is *not* appealed to in the later 1959 paper that contains his official account.) On the other hand, it cannot be right to say that modal claims should be understood as covertly made in a metalanguage whose mastery requires mastery of terms that *refer to* terms (here, descriptive ones) in an object language—which is the classical Tarski-Carnap sense. For someone (perhaps a monolingual German) could claim, believe, or judge that A causally necessitates B without ever having heard of the English *expressions* that 'A' and 'B' stand for in the example. Further, the claim could be true even if there had never been such expressions, because there had never been any language users. (There would still have been laws of nature, even if there had never been language.) So is the view he is after a metalinguistic expressivism, or not? In light of the considerations just mentioned, Sellars's characteristically nuanced-but-unhelpful assessment is this:



Shall we say that modal expressions are metalinguistic? Neither a simple ‘yes’ nor a simple ‘no’ will do.<sup>118</sup>

He wants to say that while modal statements are not metalinguistic in a narrow sense, there is a wider sense in which they are.

It is sometimes thought that modal statements do not describe states of affairs in the world, because they are *really* metalinguistic. This won’t do at all if it is meant that instead of describing states of affairs in the world, they describe linguistic habits. It is more plausible if it is meant that statements involving modal terms have the force of *prescriptive* statements about the use of certain expressions in the object language. Yet there is more than one way of to ‘*have the force of*’ a statement, and failure to distinguish between them may snowball into a serious confusion as wider implications are drawn.<sup>119</sup>

What distinction does he have in mind?

We must here, as elsewhere, draw a distinction between what we are committed to concerning the world by virtue of the fact that we have reason to make a certain assertion, and the force, in a narrower sense, of the assertion itself.<sup>120</sup>

Sellars acknowledges that modal statements do not *say that* some entailment holds, but distinguishes between what is *said* by using a bit of vocabulary and what is ‘*contextually implied*’ by doing so. Sellars says very little about this latter notion, even though it bears the full weight of his proposed emendation of the rationalist account. This is really all he says about the matter in the only essay he devotes to the exposition of his views about the “causal modalities.”

Elsewhere he had put what I think is recognizably the same point in terms of a distinction between what one *says* by making a statement and what (else) one *conveys* by doing so.<sup>121</sup> There his example is that in asserting

---

118 CDCM §82.

119 CDCM §81.

120 CDCM §101.

121 “Inference and Meaning”, p. 280/332 in in J. Sicha (ed.) *Pure Pragmatics and Possible Worlds: The Early Essays of Wilfrid Sellars* [Ridgeview Publishing Company, Reseda CA, 1980]. This is also an earlier piece (1953), and he does not in CDCM advert to this way of making the distinction.

“The weather is fine today,” I *say* that the weather is fine today, but *convey* that I *believe* that it is fine. This is suggestive, but won’t help us out in detail in the modal case. For, first, he doesn’t give us any idea what, if anything, *is said* by making a modal claim. Second, assertions are in general expressions of belief, regardless of what their content is. But the case we care about depends on the application of specifically modal concepts in what is said *doing* something specific that one is *not* doing in making assertions generally.

I think Sellars never really figures out how to work out the line of thought he suggests here. After 1959 he never repudiates the views he sketched in “Counterfactuals, Dispositions, and the Causal Modalities,” and seems to continue to endorse them. But he never revisits the topic substantially—never says how he thinks one might go on to fill in the expressivist idea he had gestured at there. Doing that is, in effect, left as an exercise to the reader. I conjecture that one reason for this failure is that he labored under the restriction of a further systematic constraint consequent upon other views near and dear to his heart. For he *also* thought that discourse about properties, universals, and even facts was metalinguistic in a broad, nonclassical sense. The problem for him, I think, is that he thought he not only needed to find a specific sense in which modal vocabulary could be understood to be ‘metalinguistic’, but also a sense of that term that was *generic* between that case and the case of ontological-categorial vocabulary such as ‘property’ and ‘universal’. He *did* work hard, and make significant progress, on delineating the sense in which he thought of that latter sort of vocabulary as metalinguistic, avoiding the pitfalls (mentioned above) involved in understanding it as metalinguistic in the orthodox sense that requires reference to the expressions of an object language. His response turns on the discursive functional roles that dot-quoted expressions refer to, the notion of distributive singular terms, and of the formation of a kind of such terms by instantiating-categorizing quotation to refer to those roles.<sup>122</sup> This is a very sophisticated response to the corresponding difficulties that arise for calling ontological-categorial expressions ‘metalinguistic’. But *that* solution does not *immediately* apply to modal expressions. (Whether some variant of it would work is another question.) And he could not figure out how to specify either the genus that comprises both, or the modal species.

Sellars is working with Kant’s idea that the expressive role distinctive of alethic modal vocabulary is to make explicit something that is implicit already in the use of ordinary empirical

---

122 His views are developed in three seminal essays: “Naming and Saying,” “Grammar and Existence: A Preface to Ontology,” and “Abstract Entities.” They are reprinted as Chapters Five, Six, and Seven of K. Scharp and R. Brandom (ed.s) *In the Space of Reasons: Selected Essays of Wilfrid Sellars* [Cambridge: Harvard University Press, 2007].

descriptive vocabulary. He picks up Frege's hint that what matters is the specifically *inferential* articulation essential to the conceptual contentfulness of descriptive vocabulary. He develops those thoughts by adding the ideas that that expressive role is in some broad but noncanonical sense metalinguistic—a matter of the role such vocabulary plays in endorsing rules of inference governing descriptive vocabulary. And equally importantly, he focuses our attention on the *pragmatic* dimension of that expressive role. That is, he counsels us to look to what we are *doing* when we endorse a modal claim. (Compare: expressivism about normative vocabulary—paradigmatically deontic vocabulary.)

I want to make a couple of suggestions for how one might move forward with what Sellars made of Kant's thought about how the expressive role characteristic of alethic modal vocabulary is related to that of ordinary empirical descriptive vocabulary. One lesson I think we can learn from Sellars's difficulties is that the notion of being 'metalinguistic' or ("about language") is too crude an expressive tool, too undifferentiated a concept, to be helpful in this context. There are, as Sellars intimates, *many* ways in which the use of one vocabulary can depend on that of another, besides any terms of the one vocabulary *referring* to those of the other. Putting together Sellars's *metalinguistic* idea with his *pragmatic* idea, we could consider the possibility that the place to begin thinking about the expressive role of modal vocabulary is with what in *Between Saying and Doing* I call a "pragmatic metavocabulary." This concept takes its place alongside that of a syntactic metavocabulary, which enables one to talk about linguistic expressions themselves (both what Sellars calls "sign designs" and grammatical categories), and a semantic metavocabulary, which enables one to talk about what linguistic expressions refer to or what descriptive concepts let one say. A *pragmatic* metavocabulary enables one to talk about what one is *doing* in *using* linguistic expressions, the speech acts one is performing, the pragmatic force one is investing them with or exercising, the commitments one is undertaking by making claims, the norms that govern linguistic performances, and so on. (This list is something of a motley, meant to correspond to the capaciousness of 'do' and 'use', a reminder that the concept is picked out is still generic.) Sellars's model is that modal vocabulary says something that would be said more explicitly in a *semantic* metavocabulary. But by the time his commentary has taken back everything that it turns out needs to be taken back, not much is left of that model. What seems right about the commentary, however, is Sellars's observations

about what one is *doing* in “making first-hand use” of modal vocabulary: endorsing inferences. Insofar as there is anything to that idea, the more natural strategy would seem to be to take one’s model from *pragmatic* metavocabularies. After all, Sellars ends up saying nothing at all about what one *says* in making first-hand use of modal vocabulary. Properly understood, I think, his is not a *semantic* expressivism about alethic modal vocabulary, but a kind of *pragmatic* expressivism about it.

As a first try at expressing the thought that would result from transposition from a semantic into a pragmatic key, we might try this: In making first-hand use of (the relevant kind of) alethic modal vocabulary one is *doing* something distinctive that could be specified explicitly in the right kind of pragmatic metavocabulary, namely endorsing a class of inferences. The pragmatic metavocabulary enables one to *say* what modal vocabulary enables one to *do*. Such a claim does not in itself involve any commitment concerning the relations between the *content* of talk about endorsing inferences and talk about necessity and possibility, never mind commitment to their equivalence. Notice, further, that counterfactuals that suppose the absence of concept users are irrelevant to the assessment of *this* claim. For in that case there would be neither endorsers of inferences nor users of modal vocabulary.

The claim that is on the table so far is evidently too weak to be interesting, though. It does not carve out an expressive role that is *distinctive* of modal vocabulary. For in making an ordinary descriptive claim one is *also* doing something that could be specified in a pragmatic metavocabulary, namely applying descriptive concepts, making a claim, undertaking a doxastic or assertional commitment. And those, the Frege-Sellars inferentialist line goes, essentially involve commitments to the proprieties of inferences. My second suggestion for developing Sellars’s modal expressivism is that what is special about (a certain kind of) modal vocabulary is that it stands in a special relation to descriptive vocabulary—a relation that invited its characterization as ‘metalinguistic’ (with respect to that descriptive vocabulary) in the first place. This relation is that anyone who knows how to use ordinary empirical descriptive vocabulary (e.g. ‘red’, ‘square’, ‘moving’, ‘alive’, ‘electron’) already knows how to do everything she needs to know how to do to deploy modal vocabulary. A variant formulation (closely related, but not equivalent) would be that the norms governing the use of ordinary empirical descriptive vocabulary determine the norms governing the use of modal vocabulary. In this sense, modal vocabulary makes

explicit (in the form of a new kind of claimable content) something that is implicit already in the *use* of descriptive vocabulary. This claim about the expressive role characteristic of modal vocabulary is vocabulary-specific. For not all vocabularies stand in this relation to some other kind of vocabulary. In particular, there is in general nothing that ordinary empirical descriptive (OED) vocabulary stands to in this expressive relation.

An instructive parallel is with a particular bit of logical vocabulary: the conditional. If Sellars is right that an essential element distinguishing *describing* from mere *labeling* keyed to differential responsiveness is the inferential involvements of the locutions applied (their “situation in a space of implications”) then anyone who knows how to use descriptive vocabulary already knows how to do everything he needs to know how to do to use conditionals whose antecedents are formed from those descriptive claimables. For to be able to use the descriptive vocabulary, one must make some distinction (however partial and fallible) between materially good and materially bad inferences involving that vocabulary. And that is sufficient to introduce conditionals as having the circumstances of appropriate application that if one is committed to the propriety of the inference from  $p$  to  $q$ , then one is committed to the conditional claim “if  $p$  then  $q$ ,” and the consequences of application that if one is committed to the conditional claim “if  $p$  then  $q$ ,” then one is committed to the material propriety of the inference from  $p$  to  $q$ . The capacity to use the underlying descriptive vocabulary can be straightforwardly (indeed, algorithmically) transformed into the capacity to use conditionals involving that vocabulary.

What aspect of inference is it that modal vocabulary is supposed to express? My third suggestion for developing the Kant-Sellars approach to modality is an answer to this question. The key fact to appreciate, I think, is that outside of logic and mathematics (and possibly fundamental physics, though I doubt it<sup>123</sup>), in ordinary language and the special sciences, material inference is massively *nonmonotonic*. That is, the fact that the inference from  $p$  to  $q$  is a materially good one in some situation does not mean that the inference from  $p$  and  $r$  to  $q$  must also be a good one, in the same situation. If I strike this dry, well-made match, it will light—but not if in addition all the oxygen is removed from the room, or a sufficiently strong magnetic field is applied, or.... If I let loose of the leash, the dog will chase the cat—but not if either one is struck by lightning, a bear suddenly blocks the way, or.... This phenomenon is ubiquitous and

---

123 For reasons Mark Wilson elaborates in his original and important book *Wandering Significance* [Oxford: Oxford University Press, 2006]

unavoidable, even in less informal contexts: differential medical diagnosis, the application of common or case law, or philosophical argumentation. One cannot secure material inferences from all possible defeasors by explicitly building their denial into the premises, for the class of defeasors is in general open-ended and not antecedently surveyable. Nor can one achieve the same effect wholesale by the use of *ceteris paribus* clauses. As I have argued elsewhere, the expressive role of such clauses is explicitly to acknowledge the non-monotonicity, hence defeasibility of the qualified inference, not magically to remove it.<sup>124</sup> (The technical term for a Latin phrase whose application can do *that* is ‘spell’).

The defeasibility or nonmonotonicity of the material inferences essential to the conceptual contentfulness of descriptive vocabulary means that the use of such vocabulary requires not only making a distinction (however fallibly) between those inferences one endorses and those one does not, but also (as part of that capacity, and also fallibly) between the collateral premises or auxiliary hypotheses whose additions one takes it would, and those that would not, infringe the inference, in the sense that the conclusion would no longer follow. That is, in order to use OED vocabulary, one must associate *some* range of subjunctive and counterfactual robustness with the material inferences that (at least partially) articulate the contents of the descriptive concepts. So, for instance, I might endorse the inference that would be made explicit in a conditional by “If I release my grip on the book, then it will fall to the floor.” But for the attribution of such an inferential commitment to me to be sustainable, I must make some distinction between collateral circumstances that would defeat the inference (a table is moved under it, someone else catches it, it dissolves in a puff of smoke, it is snatched up by a passing hawk...) and those that would not (it is Tuesday, it is slightly cooler today than it was yesterday, my car has been moved slightly further away...). Of course I might be wrong about whether any of these particular auxiliary hypotheses actually would or would not defeat the inference to the conclusion. But if I make no distinction of this sort at all I should be convicted of not understanding the concepts (book, falling) that I am attempting to apply.

---

124 In Chapter Two of *Articulating Reasons: An Introduction to Inferentialism* [Cambridge: Harvard University Press, 2000].

The principal vocabulary we use to make these distinctions explicit is *subjunctive* and *counterfactual* conditionals: “If the lioness *were* to be struck by a spear...,” “If the book *had been* attached to a large helium-filled balloon...” Subjunctives let us express, explore, and communicate the ranges of counterfactual robustness of the inferences we endorse, our commitments concerning what would and would not defeat or infirm those inferences. The subjunctive mood is a principal alethic modal construction. Talk of what is and isn’t possible or necessary if... also lets us mark out regions of monotonicity within the field of material inferences relating applications of descriptive concepts. “If the patient has a positive muscle-contraction test, it does not necessarily follow that he has malignant hyperthermia. It is possible that he has Duchesne’s dystrophy. If he has [genetic variant], then it is necessary that he has malignant hyperthermia.” “If the wood had been pressure-treated, it would not have split over the winter, but it is possible that its color would have faded.”

On this account, subjunctive robustness is the generality or “openness” Ryle found in the inferences made explicit by conditionals, and which is made explicit by modal vocabulary, including the subjunctive mood. It involves a kind of quantification over auxiliary hypotheses that would not, according to the modal claim, infirm the inference or its conclusion.<sup>125</sup> (Frege’s account of the significance of his Latin letters indicates that he agrees with Ryle.) The kind of generalization implicit in the use of subjunctive or modal vocabulary is what is invoked in *explanation*, which exhibits some conclusion as the resulting from an inference that is good as an instance of a *kind*, or in virtue of a *pattern* of good inferences. This is what was intuitively right about the deductive-nomological understanding of explanation. What was wrong about it is that subjunctive robustness need not be underwritten by *laws*: modally qualified conditionals whose quantifiers are wide open. That is, there need not be inferences guaranteed to be *globally* monotonic no matter what collateral premises are thrown in, standing behind every *local* region of monotonicity—every set of collateral premises with respect to which the inference *is* subjunctively robust. Thus singular explanations, for instance, singular causal explanations, need not fall under covering laws to be good explanations. But they do need to involve *some* range of subjunctive (including counterfactual) robustness in order to count as *explanations*, rather than just descriptions of some event. It is because the use of descriptive vocabulary requires commitment to inferences with some range of subjunctive robustness that, as I earlier quoted Sellars as saying:

---

125 Many everyday uses of modal vocabulary to qualify claims suppress the premises from which the claim implicitly is taken to follow, and so court the danger of countenancing the modal fallacy that would infer from  $p$  and  $\Box(p \rightarrow q)$  to  $\Box q$ . Thereon hangs a tale.

Although describing and explaining (predicting, retrodicting, understanding) are *distinguishable*, they are also, in an important sense, *inseparable*... The descriptive and explanatory resources of language advance hand in hand...<sup>126</sup>

The expressive job characteristic of modal vocabulary is to make explicit this implicit dimension of the use of ordinary empirical descriptive vocabulary.

## II. A Modal Realism

This sketch of a program for extending the Kant-Sellars tradition of modal expressivism raises a myriad of questions, some of detail, others more substantial. Rather than beginning to fill in that sketch by addressing some of those questions, I want to confront the ideas that motivate it with a different set of intuitions: those that motivate a robust modal realism. By “modal realism” I mean the conjunction of the claims that:

MR1) Some modally qualified claims are *true*.

MR2) Those that are state *facts*.

MR3) Some of those facts are *objective*, in the sense that they are independent of the activities of concept-users: they would be facts even if there never were or never had been concept-users.<sup>127</sup>

There are strong reasons to endorse all three of these claims. As to the first, physics tells us things such as: “Two bodies acted upon only by gravitational forces necessarily attract one another in direct proportion to the product of their masses and in inverse proportion to the square

---

126 CDCM § 108.

127 Of course, this is itself a modal claim, expressed counterfactually in the subjunctive mood. That fact is not problematic in the current context. One upshot of the previous discussion is that *any* description of how things objectively are implicitly involves modal commitments.



of the distance between their centers of mass.” I take it this claim, for instance, is true. Even if it is not, I take it that *some* claims of this form, purporting to state laws of nature, do, in fact, state laws of nature. Denying this brings one into direct contradiction with the empirical sciences themselves. Supporting such a position would require a strong argument indeed. For the empirical sciences are in the business of making subjunctive and counterfactual-supporting claims. That is, they offer not only *descriptions*, but *explanations*. Indeed, the descriptions they offer are essentially, and not just accidentally, available to figure in explanations of other descriptions.

The second claim is, I think, true in virtue of the definition of ‘fact’. A fact, Frege says, is a thought that is true.<sup>128</sup> He means ‘thought’ in the sense of something thinkable, not in the sense of a thinking, of course. For there can be unthought facts. On this usage, it is alright to say that facts make thoughts or claims true only in the sense that facts make acts of thinking and claiming true. For the facts just *are* the true thinkables and claimables. Wittgenstein is appealing to this way of using ‘fact’ when he says: “When we say, and mean, that such-and-such is the case, we—and our meaning—do not stop anywhere short of the fact; but we mean: this—is—so.”<sup>129</sup> On this usage, if there are true modal claims—in the sense of true modal claimables, or modal claimings that are true in that they are claimings of true claimables—then there are modal facts. Modal facts are just facts statable using modal vocabulary, as physical facts are facts statable using physical vocabulary, nautical facts are facts statable using nautical vocabulary, and so on.

The third claim is perhaps the most controversial of these three platitudes. But I think the same principle I implicitly invoked in talking about the first claim underwrites it. Physics tells us that the current laws of nature were already laws of nature before there were human concept-users. And, although it does not specifically address the issue, it is clearly committed to the claim that the laws would have been the same even if there never had been concept-users. Indeed, many of the laws of nature (including all the Newtonian ones) exhibit a temporal

---

128 In “The Thought” [ref.].

129 *Philosophical Investigations* [ref.] §95.

symmetry: they hold indifferently at all times. So they are independent of the advent, at some particular time, of concept-users. And one of the mainstays of physics over the last century—substantially contributing to its distinctive conceptual shape—is the result of the Noether theorem that tells us (entails) that that this fundamental temporal symmetry is mathematically equivalent to the physical principle of conservation of energy.<sup>130</sup> Denying MR3 is denying the temporal symmetry of laws of nature. And the theorem tells us that that means denying the conservation of energy. While there are reasons from the bleeding edge of physics to worry about the universal truth of the principle of conservation of energy, those considerations are irrelevant in the current context: they do not stem from the presence or absence of concept-users in our world). I conclude that one cannot deny MR3 without taking issue with substantial, indeed fundamental, empirical issues in physics.<sup>131</sup>

---

130 Cf. for instance Nina Byers (1998) "E. Noether's Discovery of the Deep Connection Between Symmetries and Conservation Laws." in *Proceedings of a Symposium on the Heritage of Emmy Noether*, held on 2–4 December, 1996, at the Bar-Ilan University, Israel.

131 I offer a different argument for this same conclusion (not specifically for the modal case, but for a more generic one that comprises it) in Section V of Chapter Five of *Perspectives on Pragmatism*.

There were no true claimings before there were vocabularies, because there were no claimings at all. But it does not follow that there were no true claimables. In fact, we can show that we ought not to say that. Here is an argument that turns on the grammatical transformations that “It is true that...” takes.

Physics tells us that there were photons before there were humans (I read a lot about them in Stephen Weinberg’s account of the early history of the universe, *The First Three Minutes* [New York: Basic Books, 1988], for instance). So if before time V there were no humans, so no vocabularies, we do not want to deny that

1.           There were (at time pre-V) photons.  
We can move the tense operator out front, and paraphrase this as:
2.           It was the case (at time pre-V) that [there are photons].  
By the basic redundancy property of ‘true’, we can preface this with “It is true that...”:
3.           It is true that [It was the case (at time pre-V) that [there are photons]].  
Now we can move the tense operator out to modify the verb in “It is true that...”:
4.           Was[ It is true (at time pre-V) that [there are photons]]

This is the key move. It is justified by the observation that *all* sentential operators can be treated this way, as a result of deep features of the redundancy of ‘true’. Thus one can transform “It is true that Not[*p*],” into Not[It is true that *p*], “It is true that Possibly[*p*],” into “Possibly[It is true that *p*],” and “It is true that Will-be[*p*],” into “Will-be[It is true that *p*].” But now, given how the tense operators work, it is straightforward to derive:

5.           It was true (at time pre-V) that [there are photons].  
And again invoking the features that make ‘true’ redundant, we get:
6.           It was the case (at time pre-V) that [It is true that [there are photons]].

These uniformities involving the interaction of ‘true’ with other sentential operators tell us we are committed by our use of those expressions to either deny that there were photons before there were people—which

I am claiming that one ought to endorse MR1 and MR3 unless one takes issue with the principle that philosophers thinking metalinguistically about semantics and concept-use ought not, in general, to be in the business of denying claims made by physicists, when the latter are speaker *ex cathedra* on matters that fall within their professional purview. There are some philosophers (Huw Price is one) who are both competent and willing to do so—indeed, in his case, specifically on the matter of the physicists’ uncritical use of modal vocabulary. But I am not one of them.

I take it that:

1) If some crucible were heated to a temperature high enough to melt copper, then it would be hot enough to melt aluminum.

Is a chemical necessity: a chemical law of nature. It is a modal fact. It is modally, subjunctively, counterfactually independent of the existence of concept-users. If that is right, then descriptions of how things objectively are stand in modally robust material (non-logical) consequential relations to one another. Another such is:

2) If the sample were (had been) pure copper, then it would be (would have been) denser than water.

Besides relations of material consequence, descriptive facts we can state can also stand in relations of material incompatibility.

3) A sample’s being pure copper is incompatible with its being an electrical insulator. (It is not possible that a sample be both pure copper and an electrical insulator.)

Ways the world can be empirically described as being stand to one another in objective, modally robust relations of material consequence and incompatibility.

The modalities this sort of realism addresses are those invoked by the natural sciences, and their analogs in less systematic ordinary language. What the kind of modal vocabulary in question expresses is not *logical* possibility and necessity, for the truth of claims such as (1), (2) and (3) depends essentially on their use of the *non-logical* empirical descriptive concepts copper, aluminum, temperature, water, density, and so on. Nor is it

---

is to deny well-entrenched deliverances of physics—or to admit that there were truths about photons before there were people to formulate them.

metaphysical necessity, which abstracts from actual laws of nature and other subjunctive- and counterfactual-supporting dependencies that turn on particular properties things can be described as having.

The modal revolution in late twentieth-century Anglophone philosophy had three principal phases. First was Kripke's revolution in the semantics of modal logical vocabulary. Second was the generalization, by Lewis, Stalnaker, Montague, and Kaplan, among others, of his algebraic possible-worlds apparatus to an intensional semantics for non-logical expressions. Third was the introduction of the conceptual apparatus that led to the recognition of the possibility of necessities knowable only *a posteriori*, and contingencies knowable *a priori*, in Kripke's *Naming and Necessity*. It was this third phase that gave rise to contemporary analytic metaphysics. The kind of modality to which both the modal expressivism of the previous section and the modal realism of this one are addressed is relevant at most to the second phase: the one in which modal notions such as possibility are used to explicate the contents of nonlogical concepts.

There is another line of argument to the conclusion that commitment to modal realism is implicit in commitment to a corresponding realism about claims expressed using ordinary empirical descriptive vocabulary. It will make clearer the relation between one kind of alethic modality and conceptual content. We can begin with a platitude: there is some way the world objectively is. How it objectively is must be discovered by empirical inquiry, and sets a semantic and epistemic standard for assessment of the correctness of our descriptive claimings as potential expressions of knowledge. The question is how to understand the relation of modal facts (if any) to how the world objectively is as describable (at least sometimes) in non-modal empirical descriptive vocabulary. One might ask a supervenience question here, but the line of thought I am concerned with goes a different way. It asks what modal commitments are implicit already in the idea of an empirically describable world. It focuses on the *determinateness* of the way things objectively are.

To talk about how things objectively are as determinate is to invoke a contrast with how they are not. This idea is summed up in the Spinozist (and scholastic) principle *omnis determinatio est negatio*. This thought is incorporated in the twentieth-century concept of

information (due to Shannon<sup>132</sup>), which understands it in terms of the partition each bit establishes between how things are (according to the information) and how they are not. But there are different ways we might follow out this idea, depending on how we think about the sort of negation involved. What I'll call the "Hegelian" model of determinateness insists that it must be understood as what he calls "exclusive" [ausschließend] difference, and not mere or "indifferent" [gleichgültig] difference.<sup>133</sup> Square and circular are exclusively different properties, since possession by a plane figure of the one excludes, rules out, or is materially incompatible with possession of the other. Square and green are merely or indifferently different, in that though they are distinct properties, possession of the one does not preclude possession of the other. An essential part of the determinate content of a property—what makes it the property it is, and not some other one—is the relations of material (non-logical) incompatibility it stands in to other determinate properties (for instance, shapes to other shapes, and colors to other colors). In fact, Hegel's view is that determinateness is a matter of standing in relations of material incompatibility (his "determinate negation") and material consequence (his "mediation") to other determinates. We might think of these as related by the principle that one property, say metallic is a consequence of another, copper, in case everything incompatible with being metallic (say, being a mammal) is incompatible with being copper. A property possession of which rules out possession of *no* other properties, and has as a consequence possession of no others, is in so far such *indeterminate*.

One observation we can make about this distinction between exclusive difference and mere difference is that one can define mere difference solely in terms of exclusive difference, but not *vice versa*. For one can say that two properties are *merely* different just in case they are not incompatible with each other, but are materially incompatible with different properties. Square and green are different because they are incompatible with different properties: square is incompatible with circular, and green is not.<sup>134</sup>

---

132 Claude E. Shannon and Warren Weaver: *The Mathematical Theory of Communication*. The University of Illinois Press, Urbana, Illinois, 1949. ISBN 0-252-72548-4

133 The rubric 'Hegelian' here is tendentious, and liable to be alarming. More seriously, it is liable to be unhelpful. For now, treat it as a mere label. I will say what I mean by it—give it some content—as we go along.

134 This definition sounds circular, because of its invocation of the notion of sameness of the properties incompatible with a property. But we can avoid this. Suppose we have labeled properties (say, by real numbers). If an oracle then tells us for each label the set of all labels of incompatible properties, we can sort the labels into

One reason to endorse this Hegelian conception of determinateness is that it is required to underwrite what might be taken to be an essential aspect of the structural difference between the fundamental ontological categories of *object* and *property*. Aristotle had already pointed out a fundamental asymmetry between these categories. It makes sense to think of each property as coming with a *converse*, in the sense of a property that is exhibited by all and only the objects that do *not* exhibit the index property. Has a mass greater than 5 grams is a property that has a converse in this sense. But it does *not* make sense to think of *objects* as coming with converses, in the analogous sense of an object that exhibits all and only the properties that are *not* exhibited by the index object. This is precisely because some of those properties will be incompatible with one another. Thus my left leg has the properties of not being identical to Bach's second Brandenburg concerto and not being identical to Gottlob Frege. Its converse, if it had one, would have to have the properties of being identical to both.

Now one might deny that this categorial asymmetry is essential to the concepts of object and property. A *Tractarian* conception of (elementary) objects and properties makes do with mere difference. Elementary properties and relations do not stand in relations of material incompatibility or consequence. They are independent, in that the fact that an object exhibits one property or stands in one relation has no consequences for any others it might exhibit or stand in.<sup>135</sup> (*All the relations of incompatibility and consequence holding between states of affairs in the Tractatus hold between non-elementary states of affairs, and are due solely to the logical complexity of those states of affairs. There are no material, that is, nonlogical, relations of consequence and incompatibility in that picture.*) In this context it is coherent to associate with each elementary object a converse, which exhibits all and only the properties (stands in all and only the relations) that the index object does not. I am not concerned here to argue that the Tractarian conception of object is incoherent or otherwise inadequate just because it has no room for the Aristotelian categorial asymmetry. For my purposes it is sufficient to point out that the Hegelian notion of determinateness, which requires acknowledging the distinction between mere difference and exclusive difference,

---

equivalence classes, accordingly as the set of incompatible labels they are associated with is the same. These will all be labels of the same property. Two labels that are not in the same incompatibility-equivalence class, are then labels of different properties. Some pairs of properties that are different in this sense will then also be exclusively different, if one is a member of the incompatibility set of (a label of) the other.

135 There are both textual and conceptual difficulties concerning the status of monadic elementary properties in the *Tractatus*. But the points I am concerned to make go through just as well if we restrict ourselves to relations, so I will ignore both these kinds of difficulty.

*does* underwrite (is necessary and sufficient for) the Aristotelian point about the difference between objects and properties (or relations).

A Tractarian conception of determinateness is one according to which it is sufficient for properties to be determinate that they are *merely* different from one another, and sufficient for objects to be determinate that they exhibit some merely different properties. Tractarian properties do not stand to one another in relations of determinable properties (e.g. polygonal, colored) and more determinate properties falling under them (circular, green). For the more determinate properties would stand in relations of material consequence to their determinables, and in relations of material incompatibility to other determinates falling under the same determinable. So nothing like the structure—characteristic of shapes and colors, and of biological taxonomies—of properties as falling into determinable families of exclusively different determinates which are merely different from determinates falling under other determinables is available in a Tractarian world.

The Hegelian conception of determinateness as a matter of standing in relations of exclusive difference (material incompatibility, and—so—material consequence) to other determinates, then, has at least these three consequences in its favor:

- The mere difference that articulates the Tractarian world can be defined in terms of exclusive difference, but there is no backwards route;
- Objects and properties that are determinate in this sense exhibit the Aristotelian categorial asymmetry;
- Properties will exhibit the standard structure of compatible determinable families of incompatible determinate properties.

It should be clear that to take the objective world to be determinate in the Hegelian sense—so, to consist of objects and their properties and relations in the Aristotelian sense, and for those properties and relations to exhibit the structure of determinable families of determinates—is to be committed to modal realism. For Hegelian determinateness requires that there be facts about what properties and states of affairs are materially incompatible with which others, and about what material consequential relations they stand in to which others. The determinateness of the fact that this coin is copper consists in part in its being incompatible with the coin being silver and its having as a consequence that it conducts electricity—that is, with its being *necessary* that

it is not silver, *possible* that it is green, and *necessary* that it conducts electricity.<sup>136</sup> Metallurgists discover these modal facts as part of the same kind of empirical inquiry through which they discover that this coin is in fact copper. A world without modal facts would be an indeterminate world: a world without objects in the Aristotelian sense, and without properties in the sense that admits a determinate-determinable structure.

The kind of modality in question is that expressed in ordinary conversational language, and in a more systematic and controlled way in the special sciences, both empirical and exact. It is the modality involved in claims such as “No monochromatic patch can be both red and green,” “It is impossible for a square plane figure to be circular,” “Pure copper at sea-level pressure necessarily melts at 1083.4 °C,” and “A mammal placed in an evacuated bell-jar would die of oxygen deprivation.” These are not either *logical* modalities, except in an extremely extended sense—though one not without precedent in Anglophone philosophy of the ‘40s and ‘50s), nor are they oomphier metaphysical modalities in a Kripkean sense.

In laying out Sellars’s views I registered that he thinks of what he called the “causal modalities” as characterizing the inferential relations that articulate the contents of empirical descriptive concepts. If we go back to what Hegel made of Kant’s views of modality and conceptual content, we find a notion of conceptual content that can help us better understand how this kind of modality can be understood as a *conceptual* modality. On this conception, to be *conceptually* contentful just is to stand in modally robust relations of material consequence and incompatibility (what Hegel calls relations of “mediation” and “determinate negation”). This is a resolutely non-psychological sense of ‘conceptual’. For it makes no reference to *concept-use*—to the *application* of concepts by anyone at all. So if there are laws of nature according to which some properties are incompatible with others (*cannot* be exemplified by the same object at the same time) or have others as their consequences (if one is exhibited by an object, the other *must* be) then the world as it is objectively, independently of the activity of any knowing and acting subjects, is conceptually articulated. Empirical inquiry is at once the job of determining what judgments are true and what concepts are correct—that is, what really follows from what and what really precludes what. Linguistic terms can *express* concepts, by being used so as to

---

136 Of course there are various provisos that would have to be added to make these claims strictly true, since copper can be alloyed with silver, and so on. I ignore these complications, as beside the point I am after.



undertake commitments as to what follows from what and what precludes what. But the concepts they express are in no sense *products* of that concept-applying activity.

As we saw, Sellars insists that it is standing in such relations that makes empirical descriptive vocabulary genuinely *descriptive*, that is, expressive of descriptive *concepts*, rather than merely functioning as reliably differentially responsively elicited *labels*. And we have seen that the sort of modal realism I have been sketching has as one of its consequences that empirical descriptive properties and states of affairs stand to one another in relations of material consequence and incompatibility. So Hegel offers us definitions of what it is to be *determinate* and to be *conceptually articulated*, according to which to take the objective world to be determinate is to take it to be *modally* articulated and to be *conceptually* articulated. That is, it commits one both to modal realism and to conceptual realism: the view that the objective world is modally, and *so* conceptually structured, quite apart from its relations to us.

### III. Together

The core of the modal realism I have just sketched consists of some claims that express philosophical common sense: there are laws of nature, events sometimes causally necessitate others, there is a determinate way the world objectively is, and its being that way rules out (excludes the possibility) of its being some other ways. These are commitments to which any philosopher ought to want to be entitled. They should be contested only under theoretical duress by exceptionally weighty and compelling arguments. I have elaborated those core claims in the context of others that are *not* commonsensical, most notably that modal realism in this sense entails *conceptual* realism about the objective world. The link between the two classes of claim is forged by the Hegelian non-psychological definition of the conceptual, as what is articulated by relations of material (that is, in general nonlogical) consequence or necessitation and incompatibility. I think this is a good thing to mean by “conceptual,” not the least because of the space it opens up to understand how the sort of causal modalities investigated by the sciences can be thought of as articulating the contents of concepts. That is a deservedly controversial claim. Whatever stance one takes on it, the sense in which I am using the term “conceptual” is, I trust, at least reasonably clear.

But what is the relation between this kind of modal *realism* and the modal *expressivism* I talked about in the first part of this essay? There the expressive role characteristic of modal vocabulary was identified as making explicit the material inferential and incompatibility relations in virtue of which ordinary empirical descriptive (OED) vocabulary expresses the content that it does. This expressive role was distinguished from that of the ground-level empirical descriptive vocabulary, whose principal job it is to say how things objectively are. There is no further vocabulary to which OED vocabulary stands in the same semantically explicative relation as alethic modal vocabulary stands to it.<sup>137</sup> The core of this version of modal expressivism lies precisely in the distinction it insists on between the expressive role distinctive of modal vocabulary and that of vocabulary whose job is describing the world, at least in the narrow, paradigmatic sense in which OED vocabulary describes the world. Modal realism says that modal vocabulary *does* describe the world, does say how things are. So are these two lines of thought simply incompatible? Are we obliged to choose between them?

I think that the modal expressivism of Part I and the modal realism of Part II are not only compatible, but that that account of the *expressive* role distinctive of modal vocabulary is just what is needed to understand the central claims of modal *realism*. The expressivism complements, rather than conflicting with, the realism about the use of modal concepts. How is such a reconciliation to be understood? The first step is to see that modal expressivism (ME) makes claims about what one is *doing* in using modal concepts, while modal realism (MR) makes claims about what one is *saying* by using modal concepts. ME says that what one is doing when one makes a modal claim is endorsing an inference relating descriptive concepts as subjunctively (including counterfactually) robust, or treating two descriptive concepts as incompatible. MR says that when one does that, onethat saying (claiming) *that* possession or exhibition of one empirical property is a consequence of, or is incompatible with, possession or exhibition of another. The claim that ME and MR are *compatible* is the claim that one can *both* be *doing* what ME says one is doing in applying modal vocabulary *and* be *saying* what MR says

---

<sup>137</sup> This is the expressive role of being *elaborated from* and *explicative of* the use of OED vocabulary. It is what in *Between Saying and Doing* I call “being LX” for that vocabulary.

one is saying by doing that. The claim that they are *complementary* is the claim that an important way to understand what one is *saying* by making modal claims is precisely to think about what one is *doing* by making them.

According to this way of understanding the relations between ME and MR, the claims of modal expressivism are made in a *pragmatic* metavocabulary for modal vocabulary: that is, a vocabulary suitable for specifying the practices, abilities, and performances that make up the *use* of modal vocabulary. And the claims of modal realism are made in a *semantic* metavocabulary for modal vocabulary: that is, a vocabulary suitable for specifying the *meanings* or conceptual *contents* expressed by modal vocabulary. What we have here is an instance of the general question of how to understand the relations between these two complementary aspects of concept application in claims: the use of the concepts and their meaning or content, what one is doing by applying them and what one is saying by applying them. I don't think we have a good general theory of how these dimensions of discourse are related to one another. (I've made a first try at an analytic framework in which such a theory might be embedded, in *Between Saying and Doing*.) Looking more closely at the special case of modal vocabulary—a vocabulary-kind of particular philosophical interest and importance—provides a potentially valuable case study and test bench for approaching the more general question of how to understand the relations between what is said in pragmatic metavocabularies and what is said in semantic metavocabularies addressing the same base vocabulary. Of special interest in this case is the relation between the use and meaning of modal vocabulary in relation to that of ordinary empirical descriptive vocabulary.

Modal expressivism says that what one is doing in making modal claims is not the same thing one is doing in making claims using ordinary empirical descriptive vocabulary. For in the former case, but not the latter, one is (perhaps *inter alia*) committing oneself to subjunctively robust inferential-and-incompatibility relations among descriptive concepts one is not in general thereby applying. Modal realism says that in making modal claims one is saying how things objectively are, describing the objective world. Reconciling these claims requires specifying a sense of “describing” or “empirical fact-stating” that is broader than that applicable to the primary use of OED vocabulary, but still sufficiently akin to it that the broader sense applicable to modal claims and the narrower sense applicable show up as species of a recognizably descriptive genus.

One broader sense that is available is that provided by *declarativism* about description, which makes it equivalent to "fact-stating" in a very capacious sense. This is the view that identifies facts with whatever is stated by declarative sentences that can be used both free-standing, to make assertions, and in embedded contexts, paradigmatically as the antecedent of conditionals and in the context of propositional attitude ascribing locutions. I think this is a perfectly good way to use "fact" and "fact-stating." But in this context, it buys modal realism too cheaply, and hence buys too cheap a version of modal realism. For in this sense "One ought not to be cruel," "Raspberries are preferable to strawberries," and "The value of Picasso's *Guernica* does not lie in its beauty," are all straightforwardly fact-stating (if they were true, they would state facts), and hence descriptive in the declarativists *very* broad sense. So this usage loses the contrast between description and evaluation (which perhaps is no bad thing, but should be a position reached for more specific reasons than the broad charter of declarativism offers) and between objective description and subjective expression of preference or other attitude. A modal realism worthy of the name should be held to a more demanding standard for what counts as empirical fact-stating or description. I conclude that a proper reconciliation of ME and MR requires crafting a sense of "empirical description" or "empirical fact-stating" that is wider than the narrow senses applicable only to OED vocabulary such as "cat", "red", and "mass of five grams", but not as broad as the declarativist's.<sup>138</sup>

Before indicating how that might be done, I want to consider one way in which the modal expressivist line of thought can be seen to be essential to understanding the modal realist line of thought. Modal realism claims that there are objective modal *facts*. One important species of modal facts is *laws* of nature. Modal realism makes essential use of the concepts of fact and law, but does not by itself *explain* those concepts. Modal expressivism does. As I indicated at the beginning of Part II, facts are (at least) true claimables. (The problem with declarativism is not its acknowledgement of this as a *necessary* condition on facts, but with its insouciant commitment to this being also a *sufficient* condition. We'll see in (10) what more might be demanded, at least for objective empirical facts.) Does this mean that there are no facts that cannot be stated, that is, expressed in some language or vocabulary? I think we adequately acknowledge the intuitive language-transcendence of fact by affirming that for any vocabulary, there are facts that cannot be stated in that vocabulary. I think of this claim as a commitment, should you specify a vocabulary, to being able to find some facts not statable in it. (I don't think, for instance, that one can express in the language of physics facts such as that the stock market dropped yesterday, or that the Republicans' unwillingness to allow a vote on the judicial nominee was a strategic political blunder.) But I don't know how to understand a claim that reverses the quantifiers and asserts that there are facts such that no vocabulary

---

138 Here I've run back and forth indiscriminately between description (or fact-stating) and empirical description as the concept being considered. I think it is the combination that matters for modal realism. These issues will be taken up separately in sections (9) and (10).

can state them. It might well be possible to give some sense to this sort of wide open quantification over all possible vocabularies, but it does not already come with one.

More deeply, though, the claim is that key concepts of the *semantic* metavocabulary in which modal realism is stated are *sense-dependent* on concepts drawn from the *pragmatic* metavocabulary for modality offered by modal expressivism. One cannot understand the concepts fact and law except in a context that includes the concepts asserting and inferring. For facts are essentially, and not just accidentally, something that can be asserted. If one does not know that it is at least sometimes true to that facts can be stated, one does not know what facts are. And laws are essentially, and not just accidentally, something that support subjunctively and counterfactually robust inferences. If one does not understand that Newton's second law of motion implies that if a force *were* (had been) applied to this moving body, it would accelerate (have accelerated), one does not grasp "F=ma" as having the force of a law.<sup>139</sup>

One concept is sense-dependent on another if one cannot grasp or understand the first without grasping or understanding the second. This sense-dependence relation must not be confused with that of reference-dependence of one concept on another, which holds when the first cannot be true of something unless the second is true of something. The concepts parent and child are both reciprocally sense-dependent and reciprocally reference-dependent. One cannot understand one in isolation from an understanding of the other, and nothing can be a parent unless something is a child (indeed, it's child), and *vice versa*. But there can be sense-dependence relations without corresponding reference-dependence relations. This is true of response-dependent properties. Suppose we define something as *hedonically beautiful for humans* just in case a human observer would respond to its perceptible presence with a feeling of pleasure. One cannot understand this dispositional property without also understanding the concept of pleasure (and, indeed, of human). But the exhibition of this property by an object does not require that there actually be feelings of pleasure. We can make perfect sense of the claim that there were sunsets that were hedonically beautiful for humans before there were humans. For to say that is just to say that *if there had been* humans to perceive them, those sunsets *would have* produced feelings of pleasure. And that can be true in a world without humans or pleasure. Similarly, if we define a planet as *supraterrestrial* just in case it has a mass larger than

---

139 In articles such as "Abstract Entities" and "Grammar and Existence: A Preface to Ontology" (reprinted as Chapters 7 and 6 of Kevin Scharp and Robert Brandom (eds.) *In the Space of Reasons: Selected Papers of Wilfrid Sellars*, [Harvard University Press 2008] Sellars develops what he calls a "metalinguistic" approach to ontological-categorial concepts such as fact and property, which is much better worked-out than his corresponding views on modality. Here, too, I think his basically Carnapian concept of the metalinguistic is far too undifferentiated to do the work he needs it to do in order to express the insights by which he is motivated. I discuss his pragmatic expressive nominalism in Chapter Seven.

that of the Earth, that concept is sense-dependent on that of the Earth, but we can use it to describe a possible world in which *all* planets are supraterrrestrial, and the Earth does not exist.

To claim that the concepts fact and law were reference-dependent on the concepts of asserting and inferring would be to assert an objectionable and obviously false sort of language- or mind-dependence of crucial categorical features of the objective world. But to claim the corresponding sense-dependence claim is not in the same way objectionable. For it is compatible with the truth of the counterfactual that there would have been facts and laws even if there had never been asserters and inferers—indeed that in our world there were facts and laws before there were asserters and inferers. The claim is just that one cannot understand what one is saying when one talks about an objective world characterized by facts and laws (which is to say just a determinate world) unless one understands facts as the kind of thing that can be stated and laws as the kind of thing that can support subjunctively and counterfactually robust reasoning. Modal expressivism helps explain what the claims of modal realism mean.

Modal realism asserts that modal vocabulary is used to form empirical descriptions of objective facts. Modal expressivism asserts that modal vocabulary plays a content-explicating expressive role that distinguishes it sharply from that of ordinary empirical descriptive vocabulary. Saying something about the broader sense in which modal vocabulary can nonetheless be understood as descriptive will further illuminate the complex complementary relations between what MR says about modal vocabulary in a semantic metavocabulary and what ME says about it in a pragmatic one. Here is a suggestion: A broader sense of “fact-stating” and “description” that is not yet so promiscuous as the declarativist candidate is defined by the dual requirements of *semantic government* of claimings by facts and *epistemic tracking* of facts by claimings.

By “semantic government” I mean that descriptive claims are subject to a distinctive kind of ought-to-be (related only in complicated ways to the ought-to-dos that Sellars contrasted them with). It ought to be the case that the content of a descriptive claiming stands in a special relation, which we might as well call “correspondence,” to a modal fact, which it accordingly purports to state (and in case there is such a fact, succeeds in stating). In virtue of that semantic norm, claimings are answerable for their correctness (accord with that norm) to facts. The underlying thought here is that what one is talking *about* is what exercises a certain kind of *authority* over what one says; what one says is *responsible to* what one is talking about, in a way that is characteristic of this relation as *semantic*. What one is talking about provides a standard for the assessment of what one says.

What is the nature of the correspondence that the norm enjoins? The contents of possible claimings are articulated by relations of material consequence and incompatibility to the contents of other potential claimings. These notions are themselves specifiable in a *deontic normative* pragmatic metavocabulary: committing (or entitling) oneself to one claim can commit (or entitle) one to others, and can preclude entitlement to still others. The contents of facts and possible facts are also articulated by relations of material consequence and incompatibility to the contents of other possible facts. In this case, these notions are specifiable in an *alethic modal* semantic metavocabulary: the obtaining of one fact has the obtaining of others as a necessary (that is, subjunctively, including counterfactually, robust) consequence, makes others possible, and rules out still others as not possible. Normative semantic government of claimings by facts says that it ought to be the case that there is a fact whose content is articulated by objective modal relations of material consequence and incompatibility that line up with the subjective (in the sense of pertaining to knowing and acting discursive subjects) normative relations of material consequence and incompatibility that articulate the content of a claiming. If that norm is not satisfied, the claiming does not live up to the standard provided by the fact it purports to state.<sup>140</sup>

Where semantic government of claiming by facts is a (deontic) *normative* matter, epistemic tracking of facts by claimings is a(n) (alethic) *modal* one. It is a matter of the subjunctive and counterfactual robustness of the conceptual content correspondence between facts and claims. The tracking condition holds just insofar as the subjunctive conditional “If the fact were (or had been) different, the claiming would be (or would have been) correspondingly different,” is true. Insofar as this condition holds, there is a *reliable* correspondence between the contents of facts and the contents of claimings. That is to say that the inference from a claim about the content of a claiming to the content of the corresponding fact is in general a good one. [I have written elsewhere about the sense in which deontic normative and alethic modal vocabularies are two sides of one (intentional) coin. I cannot here pursue this significance of this particular application (to the complementary

---

140 The concept of propositional content as what is articulated by relations of material consequence and incompatibility is a development of the Fregean metaconceptual semantic dimension of *Sinn*, while the normative relation of aboutness between objective facts and subjective commitments is a development of his metaconceptual semantic dimension of Bedeutung.

conditions of semantic governance and epistemically tracking) of that general (meta-)conceptual complementarity.<sup>141]</sup>

I think it is a fundamental mistake to try to do all the work of done by the concept of semantic government with that of epistemic tracking, as for instance Fodor and Dretske do. What goes missing is the fine structure of the crucial interaction between activities on the part of the claiming subject, expressed in a deontic normative pragmatic metavocabulary, and how it is with the objects and facts those claims are about, expressed in an alethic modal semantic metavocabulary, and how the two sides stand in both normative relations of semantic government and modal relations of epistemic tracking. It is precisely in these intricate relations that the complementary character of modal expressivism and modal realism becomes visible.

When the two requirements of semantic government and epistemic tracking are satisfied, it makes good sense to think of the claimings in question as fact-stating and descriptive. They purport to say how things are with what they in the normative sense of semantic government *about*. The actual applications of the vocabulary in question, no less than their normative status as correct or not, are epistemically *responsive* to and *controlled* by the corresponding facts. The notions of correspondence, semantic government, and epistemic tracking do not invoke causal connection—only subjunctively robust reliable covariation. For this reason, they define a notion of description or fact-stating that applies equally well to mathematical vocabulary as to empirical.

This is also evidently true also of modal vocabulary, supposing we grant the dual claims of modal realism and modal expressivism. For modal *expressivism* tells us that modal vocabulary makes explicit normatively significant relations of subjunctively robust material consequence and incompatibility among claimable (hence propositional) contents in virtue of which ordinary empirical descriptive vocabulary *describes* and does not merely *label*, *discriminate*, or *classify*. And modal *realism* tells us that there are modal facts, concerning the subjunctively robust relations of material consequence and incompatibility in virtue of which

---

<sup>141</sup> For instance, in Chapter Six of *Between Saying and Doing*.



ordinary empirical descriptive properties and facts are determinate. Together, these two claims give a definite sense to the possibility of the correspondence of modal claimings with modal facts. If we can then say what it is for a norm of semantic governance to be instituted and the modal fact of epistemic tracking to be achieved, the descriptive, the fact-stating character of modal vocabulary according to ME and MR will have been made intelligible.

It is a consequence of the version of Kant-Sellars modal expressivism that I outlined in Part I that instituting normative semantic government of modal claims by modal facts, and of achieving modal epistemic tracking of modal facts by modal claims must be an aspect of the process of instituting semantic government of ordinary empirical descriptive claims by the facts they state, and of achieving epistemic tracking of those facts by ordinary empirical descriptive claims. For the essence of that view is that what is expressed explicitly (that is, put in claimable, propositional form) by the use of modal vocabulary is already implicit in the norms governing the use of OED vocabulary.

Empiricism, in both its traditional and its twentieth century logical forms, offered a three-stage layercake picture of empirical inquiry that is particularly clear in Carnap's version. The task of the first stage is semantic: to determine the empirical concepts to be used, to fix the meanings to be expressed by OED vocabulary. The task of the second stage is epistemic: to settle, on the basis of the meanings fixed at the first stage, the claims expressed using that vocabulary that are taken to be true. The task of the third stage is explanatory: to identify, on the basis of regularities exhibited by the claims made at the second stage, laws governing the facts stated at the second stage. The first stage is a matter of convenient conventions, the last two of the assessment of empirical evidence—fraught at the second stage by the potentially problematic transition from applying observational descriptive vocabulary to applying theoretical descriptive vocabulary, and at the second stage by the potentially problematic transition from observed regularity to conjectured law. Quine sees that separating the first two stages, which makes good sense when one's model is artificial languages, is not possible when one addresses natural languages. There is just one thing discursive practitioners do: use vocabulary to make claims. Doing that must be understood as at once fixing meanings and beliefs, language and theory. Like Hume, Quine doesn't think the third stage can be rationally warranted—though this empiricist conclusion sits ill with his avowed scientific naturalism. But modal expressivism is motivated by the same pragmatic considerations about the use of vocabularies that motivate Quine's recognition of the semantic and epistemic enterprises as aspects of one process of empirical inquiry. As Sellars puts the point (in a passage I quote at the end of section 3): “although describing and explaining...are *distinguishable*, they are

also, in an important sense, *inseparable*... the descriptive and explanatory resources of the language advance hand in hand.”

Determining and applying descriptive concepts inevitably involves committing oneself as to the subjunctively robust inferential and incompatibility relations they stand in to one another. Rectifying concepts, determining facts, and establishing laws are all projects that must be pursued together. Empirical evidence bears on all of the semantic, epistemic, and explanatory tasks at once, or it bears on none of them. Of course, there is a lot more that needs to be said about how this actually works and should work. The multifarious ways in which commitments of one sort—semantic, doxastic, subjunctive—bear on and can be traded off for commitments of other sorts needs to be investigated and explicated in detail. (I’ve sketched a story about the next level of gross structure in the first three chapters of *Reason in Philosophy*.) And I certainly would not claim that seeing how modal expressivism and modal realism complement and illuminate one another clears up at a stroke all the vexing problems in the epistemology of modality—even when pursued outside the confines of the straitjacket of empiricism. But all I need here is the general conclusion—which gives us confidence that there must be solutions to those problems.

If that is right, then modal claims (and the concepts that articulate them) exhibit semantic government by and epistemic tracking of facts no less than ordinary empirical descriptive ones do. Far from being incompatible with this fundamental modally realistic claim, modal expressivism is just what is needed to make it intelligible. By showing how the use of modal concepts and the use of ordinary empirical descriptive concepts are inextricably bound up with one another, modal expressivism also shows itself and modal realism as two sides of one coin.

#### **IV. Again**

I have argued that modal realism and the right kind of modal expressivism belong together. The tendency to understand views of this kind as incompatible alternatives—to take the sense in which modal vocabulary plays, as Sellars put it a “metalinguistic” expressive role relative to ordinary empirical descriptive vocabulary to rule out the possibility of its being also fact-stating and descriptive of something other than language use—is the result of failing to attend to the

distinction between *pragmatic* and *semantic* metavocabularies. I think we don't know very much about the various ways in which what is said in these two sorts of metavocabulary are related for various vocabularies they might address. In *Between Saying and Doing*, I explore the expressive roles of various kinds of pragmatically mediated semantic relations between vocabularies, a genus that includes pragmatic metavocabularies, without saying much at all about the relations between what they make explicit and what is made explicit by traditional semantic metavocabularies of the Tarski-Carnap variety. (This was the only model Sellars had available, Procrustean though it made his efforts to formulate what I take to be his pragmatic expressivist insights.) One of my aspirations in the present piece has been to begin the process of investigating those crucial relations by looking as a test-case at a vocabulary of particular philosophical interest and importance: alethic modal vocabulary. I hope the results will be of interest to those moved by expressivist intuitions concerning other vocabularies: some kinds of normative vocabulary, moral or aesthetic, for instance, or even (were we to follow Sellars in his metalinguistic nominalism about universals) ontological-categorial or metaphysical vocabularies.

I have finished my argument. But I want to close with a lagniappe, indicated in the final word of my title. Why claim, as that title does, that the result of this story is to put modal expressivism and modal realism together *again*? Why should the story be thought of as recounting a *reunion*? The answer I want to leave you with is this: It is because we've seen something very like this constellation of metaconceptual commitments before. I started my story with Kant, and that is where I want to end it. Claiming that one should be a *pragmatic* modal expressivist (an expressivist about what one is *doing* in applying modal vocabulary) but a *semantic* modal realist (a realist about what one is *saying* in applying modal vocabulary) is, I think, recognizably a development and a descendant, for this special but central case, of Kant's claim that one should be a *transcendental idealist*, but an *empirical realist*. That is what I mean by saying that the view I have been presenting puts modal expressivism and modal realism together *again*.

## 2014 Nordic Pragmatism Lectures:

### Analytic Pragmatism, Expressivism, and Modality

#### Lecture 6

#### Sellars's Metalinguistic Expressivist Nominalism

##### I. Introduction

The five years from 1958 through 1962 were extraordinarily productive ones for Wilfrid Sellars. His monumental “Counterfactuals, Dispositions, and the Causal Modalities,” appearing in 1958, was a suitable follow-up to “Empiricism and the Philosophy of Mind” (which had been delivered as three lectures at the University of London in 1956).<sup>142</sup> Sellars never further developed the expressivist approach to alethic modality that he sketched in that paper, apparently having taken the ideas there as far as he could.<sup>143</sup> In that same year, he delivered two lectures at Yale, under the title “Grammar and Existence: A Preface to Ontology,” (*GE*, 1958) which announced an expressivist, nominalist project in ontology that he then pursued in two other equally remarkable and original essays: “Naming and Saying” (*NS*, 1962) and “Abstract

---

<sup>142</sup> *CDCM* appeared in *Minnesota Studies in the Philosophy of Science, Volume II: Concepts, Theories, and the Mind-Body Problem*, ed. Herbert Feigl, Michael Scriven, and Grover Maxwell (Minneapolis: University of Minnesota Press, 1958), p.225-308. *EPM* is reprinted in *Empiricism and the Philosophy of Mind*, Robert B. Brandom (ed.) [Harvard University Press, 1997].

<sup>143</sup> I assess how far he got, and speculate about the difficulties that could have prevented further progress, in “Modal Expressivism and Modal Realism, Together Again,” (henceforth *MEMRTA*) forthcoming.

Entities” (*AE*, 1963).<sup>144</sup> Jumblese, dot-quotes, and distributive singular terms, the conceptual tools he developed and deployed in those essays to respond to the challenges to his approach to universals he had identified in *GE*, were to remain at the center of Sellars’s philosophical enterprise for the rest of his life. Taken as a whole, these three essays provide an unusually detailed picture of the philosophical process through which Sellars progressed from an initial characterization of problems whose solutions he could not see clearly to the introduction of novel conceptual machinery that solved those problems to his durable satisfaction.

Sellars’s point of departure is a view Carnap had put forward in *The Logical Syntax of Language*: to say that triangularity is a property is a way of saying in the material mode (the object language) what is said more perspicuously in the formal mode (in a certain kind of metalanguage) as “‘triangular’ is a monadic predicate.”<sup>145</sup> This is the idea he is committed to making work in the three essays on nominalism. What Sellars calls “classifying contexts” are uses of ontological-categorical vocabulary, paradigmatically common nouns for ontological categories such as ‘property’ and ‘kind’ (and their genus, ‘universal’), the property and kind names that fall under such common nouns (‘triangularity’, ‘lionhood’), and the higher-order relations those properties and kinds are taken to stand in to their instances (such as ‘exemplification’ in “Anything that is triangular exemplifies triangularity.”). The Carnapian idea is that vocabulary of these sorts is covertly *metalinguistic*. Its use appears to tell us something about the world: what kinds (ontological categories) of things are in it. There are not only particulars, but also their properties and kinds, related to those particulars by the distinctive relation of exemplification. But actually, the claim is, the information conveyed by the use of such ontological vocabulary concerns the syntactic form of language or thought, and is not about the world talked or thought about. “Lionhood is a kind,” really means “‘Lion’ is a common noun (sortal expression).”

Such a metalinguistic species of nominalism about universals would have obvious attractions to those already of a nominalistic bent (perhaps due to a taste for desert landscapes). Is there any reason that those not

---

<sup>144</sup> All three essays are reprinted in *In the Space of Reasons: Selected Essays of Wilfrid Sellars*, Kevin Scharp and Robert Brandom (eds.) [Harvard University Press, 2007].

<sup>145</sup> Like Sellars, I will use “triangular” as short for “...is triangular”, where confusion is not likely to result.

already hagridden by nominalistic commitments should take it seriously? One potentially powerful argument is that anyone who knows how to use predicates such as “...is triangular” or common nouns such as “lion” already knows how to do everything they need to know how to do to use abstract terms such as ‘triangular’ and ‘lionhood’, and categorizing vocabulary such as ‘property’ and ‘kind’. Sellars says:

[T]o know how to use singular terms ending in '-ity' is to know that they are formed from adjectives; while to know how to use the common noun 'quality' is (roughly) to know that its well-formed singular sentences are of the form ‘— is a quality’ where the blank is appropriately filled by an abstract noun. (That the parallel points about '-keit' and 'Qualität' in German are genuine parallels is clear.)

Thus, while my ability to use ‘triangular’ understandingly involves an ability to use sentences of the form ‘— is triangular’ in reporting and describing matters of physical, extralinguistic fact, my ability to use ‘triangularity’ understandingly involves no new dimension of the reporting and describing of extralinguistic fact—no scrutiny of abstract entities—but constitutes, rather, my grasp of the adjectival role of ‘triangular’.<sup>146</sup>

‘Triangularity’ and ‘lionhood’ are singular terms formed by nominalizing adjectives and sortal common nouns, and ‘property’, ‘quality’, and ‘kind’ are categorizing sortals under which those nominalized adjectives and common nouns fall. Of course this consideration is not immediately decisive, since we can imagine a Bergmannian language in which one first learned to respond to triangular things by applying “...exemplifies triangularity,” and only later, on that basis, learned to use “...is triangular.” Nonetheless, it seems clear that one must begin by using expressions that are equivalent to predicates (adjectives): ground-level classifications. Even in the Bergmannian context, higher-order ontological classifiers such as ‘property’ will still be sortals that apply to nominalizations of these.

In *GE*, Sellars identifies two major objections that any metalinguistic nominalism about properties and kinds (universals) of this shape must face. The first is that ontologically categorizing statements such as “Triangularity is a property,” do not mention linguistic expressions, while their proposed paraphrases, such as “‘Triangular’ is a monadic predicate,” do. This difference becomes clear when we think about translating both the ontologically categorizing sentence and its explicitly syntactic paraphrase into another language. “‘Triangular’ ist ein Prädikat,” and “‘Dreieckig’ ist ein Prädikat,” are not equivalent. Which one is supposed to be the correct paraphrase of “Dreieckigkeit ist eine Eigenschaft,” which translates “Triangularity is a property”? The difference between the material mode statement and its

---

<sup>146</sup> *GE* Section XIV.

supposed paraphrase into the formal mode is even more striking when we consider counterfactuals involving them. Presumably, “Triangularity is a property” would still have been true even if the English language had never developed. Not so “‘Triangular’ is a predicate.”<sup>147</sup> If the claim that “‘Triangularity’ is a property,” is “covertly metalinguistic” or “quasi-syntactic” in character is to be sustainable in the face of these facts, the qualifications “covertly” and “quasi-” will have to be explicated in a way that avoids these consequences.<sup>148</sup>

The second objection Sellars considers is, in effect, that metalinguistic nominalism would be at best a half-hearted nominalism. For it does not avoid ontological commitment to properties (or universals, more generally). Rather, it eliminates nonlinguistic properties and kinds for linguistic ones. In place of *triangularity* and *lionhood* we get *predicatehood*, and *sortalhood*, the kinds to which belong everything that has the property of *being a predicate* or *being a sortal*. It seems that metalinguistic nominalism cannot do without expression-kinds and properties of linguistic expressions. Unlike the previous objection, this one does not directly address the adequacy of a metalinguistic account of the expressive role of ontological classifying vocabulary. It just points out that such an account is only *locally* deflationary about property-talk and kind-talk, remaining committed to it as regards *linguistic* properties and kinds.

In the large, the project Sellars announces in “Grammar and Existence”, motivates in “Naming and Saying,” and completes in “Abstract Entities” is to refine Carnap’s deflationary, expressivist idea that ontological category vocabulary is fundamentally metalinguistic, by developing it in a way that is immune to these two fundamental objections. In what follows, I describe how he does that, and critically assess the result. In brief, his response to the first objection is to introduce the technical apparatus of dot quotation, formed according to what Sellars calls the “illustrating sign-design principle.” His response to the second is to introduce further technical apparatus: the notion of distributive singular terms. This linguistic device plays a central role in drawing a distinction between what could be called “two grades of

---

<sup>147</sup> Cf. Sellars’s Section XIV of *GE*.

<sup>148</sup> “Quasi-syntactic” is the technical term Carnap uses in *The Logical Syntax of Language* for material mode expressions that should be given metalinguistic analyses.

nominalistic involvement.” Sellars distinguishes a broader notion of repeatability from a notion of universality, under the slogan “the problem of ‘the one and the many’ is broader than the problem of universals.”<sup>149</sup> He designs his metalinguistic nominalism so that the linguistic repeatables that replace worldly universals in his theory are not universals in the narrow sense.

The main critical claim I want to defend is in three parts. First, Sellars’s subtle and sophisticated development of Carnap’s metalinguistic nominalism in fact gives us a good account of the expressive role characteristic of the vocabulary of ontological categories, in particular of terms such as ‘triangularity’, ‘lionhood’, ‘property’, and ‘kind’. Second, though, I want to claim that he misunderstands the significance of this penetrating analysis. What he offers is best understood as an account of what speakers are *doing* when they say things like “‘Triangularity’ is a property,” namely, classifying expressions that play the same conceptual role as the English “...is triangular” and the German “...ist dreieckig” as adjectives. The nominalistic conclusion he wants to support, however, concerns not what one is *doing* in saying “‘Triangularity’ is a property,” but what one is *saying* by doing that. His analysis is properly understood as conducted in a *pragmatic* metavocabulary, but the conclusions he draws must be formulated in a *semantic* metavocabulary. Lacking the concept of a pragmatic metavocabulary, Sellars is not in a position to separate these considerations. Sellars’s analysis is *compatible* with semantic nominalism about universals, but does not provide an *argument* for it.

Third, I discuss the largely independent motivation for nominalism about universals that Sellars offers in “Naming and Saying.” This is epitomized in his introduction of a third bit of original technical apparatus: the language Jumblese. This argument, too, turns on the transition from a fundamental pragmatic observation about the *use* of language—that *predicating* is a kind of doing that is in principle only intelligible in terms of *saying* (asserting) and *naming* (referring), which are accordingly more conceptually basic kinds of discursive doing—to controversial claims about semantics and ontology. Its essential reliance on inferences of these forms, from what one is *doing* to what one is *saying* by doing that, shows Sellars’s *metalinguistic* semantic and ontological *nominalism* to be a particular kind of *pragmatist expressivism*.

## II. Dot Quotes and the Objection from Language Relativity

---

<sup>149</sup> *AE* p. 166.



The divergent behavior of “Triangularity is a property,” and “‘...is triangular’ is an adjective,” under translation and in various counterfactual circumstances shows that ontologically categorizing vocabulary such as ‘property’ and property-terms such as ‘triangularity’ are not metalinguistic in the narrow sense (Tarski’s) of being common nouns and singular terms falling under them that refer to the expressions of a particular object-language, such as English. This does not mean that they could not be understood to be metalinguistic in a broader sense. To specify such a sense, Sellars introduces the idea of a special kind of quotation: dot-quotation. Generically, like other forms of quotation, it is a mechanism for forming expressions from expressions. It does not, however, form *names* of expressions. Indeed, it does not form singular terms at all. I have the impression that many readers of Sellars think of dot-quoted expressions as being names of functional or conceptual roles: that •triangular• names the conceptual role played by ‘triangular’ in English.<sup>150</sup> This is not right, and in the context of Sellars’s version of nominalism about properties, it is absolutely essential to see why it is not right.

The principal features of expressions formed using dot-quotes are:

1. All expressions formed by dot-quoting other expressions are common nouns (sortals), not singular terms. That is why their basic use is in conjunction with indefinite articles as in “‘dreieckig’ is a •triangular•,” (compare: “Rex is a dog,”) or, equivalently “‘dreieckig’s are •triangular•s” (compare: “Terriers are dogs”).
2. The items falling under this kind of common noun are expression-types.
3. All the items falling under a particular common noun formed by dot-quoting an expression stand to the type of that expression in the equivalence relation ...plays the same functional-conceptual role as\_\_\_\_.

So if  $e$  and  $e'$  are specifications of expression-types,  $e'$  is a • $e$ • just in case  $e'$  plays the same conceptual role in *its* language that  $e$  plays in its language. Because ...plays the same functional-

---

<sup>150</sup> I blush to confess that I have spoken and even written carelessly in this way myself—but even Sellars himself is not always as careful on this point as he teaches us to be in *AE*.

conceptual role as \_\_\_ is an equivalence relation, one *could* treat it as an abstractor, and appeal to it to define an abstract singular term that *does* refer to the conceptual role shared by all the expression-types that stand in that relation to one another. (Perhaps one thinks of it as a name of the equivalence class defined by that relation—though that construal is certainly not obligatory.) But that is not what dot-quotes do. They would not be of much help to a program of working out a deflationary nominalist analysis of abstract entities such as properties if they did. They do serve a broadly *classificatory* function, producing a common noun that applies to all the expressions that share a conceptual role. But they do not do so by abstraction. This distinction, and the possibility it enforces of classifying without abstracting, is central to Sellars’s response to the second objection to metalinguistic nominalism.

Sellars is rather casual about the equivalence relation other expression-types must stand in to the type of the illustrating expression in order to fall under the common noun that results from dot-quoting it. He talks indifferently about “playing the same role,” “serving the same function,” “performing the same office,” and “doing the same job.” He is happy to call it a “functional” role, or a “conceptual” role. He says that what is at issue is the *prescriptive* relations it stands in to other expressions, not the descriptive ones, so he is clearly thinking about roles articulated in *normative* terms. He explicates this point by analogy to the role played by the pawn in chess. In a footnote, he indicates that he thinks these roles can be specified in terms of (norms governing) the language-entry, language-language, and language-exit transitions of a language.<sup>151</sup> I think Sellars’s lack of specificity here should be seen as evidence that the relation ... (in English) functions similarly to \_\_\_ (in German) should be seen as a placeholder, or parameter. Filling in the respects of similarity in some definite way gives rise to a correspondingly definite specification of the meaning of a particular dot-quoting locution. Dot-quoting is intended to be a *kind* of quotation, comprising as many species as there are respects of similarity of function. The elasticity of the notion of prescriptive features of conceptual or functional role should be regarded as a feature of the account, not an oversight in it.

The expression-token that appears between dot-quotes specifies the class of role-equivalent expression-types that fall under the sortal formed by the dot-quotes by *illustrating* it. The class in question is all the expression-types that are role-equivalent to the type of the quoted token. This is the “illustrating sign-design principle.” This is a kind of use of the quoted expression that is more than a mere mention of it. For, unlike standard quotation, which does

---

<sup>151</sup> Pp. 176-179. The footnote in question is #13.

merely mention the quoted expression, one cannot *understand* something of the form  $\bullet e \bullet$  unless one understands the quoted expression  $e$ . For unless one grasps the conceptual role  $e$  plays in its home language, one does not know how to tell what other expression-types stand to it in the ...plays the same functional-conceptual role as \_\_\_ relation, and so does not know what expression-types fall under the sortal  $\bullet e \bullet$ .

Expressions formed using dot-quotes are metalinguistic in a straightforward sense. They are common nouns that apply to expression-types. Sellars's idea for developing Carnap's metalinguistic analysis of what appear on the surface to be names of properties or universals, like 'triangularity' and 'lionhood', is to analyze them semantically in terms of this sort of common noun. Ontologically classifying contexts, such as "Triangularity is a property," and "Lionhood is a kind," he analyzes as " $\bullet \dots$  is triangular  $\bullet$ s are adjectives," and " $\bullet$ lion  $\bullet$ s are common nouns." *This* kind of metalinguistic statement is not subject to the first objection to Carnap's simpler version. Though they are statements in English (extended by adding some technical apparatus), they do not *refer* specifically to expressions of any particular language. Unlike ordinary quotation, but like "Triangularity is a property," and "Lionhood is a kind," they can be translated into other languages. The illustrating expressions, from which the dot-quotes are formed, can be translated right along with the rest of the sentences in which they are used. And just as it is true that even if there had never been English speakers, triangularity would still have been a property, it is true that even if there had never been English speakers,  $\bullet \dots$  is triangular  $\bullet$ s would still have been adjectives." (To deal with counterfactuals regarding the absence of language altogether, we must allow the expression-types that fall under common nouns formed by dot-quotation to include virtual ones, that is, expression-types in merely possible languages.) I conclude that the apparatus of dot-quotation permits Sellars to formulate a successor-theory to Carnap's that retains the motivating strategy of metalinguistic analysis, while successfully immunizing itself against the first objection.

### III. Two Kinds of Repeatables, Two Grades of Abstract Involvement

Addressing the second principal objection to the claim that abstract entity talk is metalinguistic, requires more than the crafting of a sophisticated extended sense of ‘metalinguistic’ (epitomized by the technical notion of dot-quotation), however.<sup>152</sup> It requires thinking hard about the nature and motivation of nominalistic commitments concerning abstract entities. For understanding triangularity in terms of •triangular•s—as in the formulation “To say that triangularity is a property is to say that •triangular•s are monadic predicates,”—is understanding the candidate abstract entity triangularity in terms of the linguistic expression-*type* •triangular•. And expression-types are themselves repeatables, under which various possible expression tokenings (in different actual and possible languages) can fall. So it would seem that being a •triangular• is a property that expressions (for instance, “dreieckig” in German) can have. In that case, nonlinguistic abstract entities, such as the property of triangularity (which triangular things have), are being analyzed in terms of linguistic abstract entities, such as the property of being a •triangular•. That suggests that metalinguistic nominalism about abstract entities is only a half-hearted nominalism, rejecting, it seems, only nonlinguistic abstract entities, but embracing linguistic ones. Such a view would in turn raise the question of the motivation for such a metalinguistic form of nominalism. Why should it be seen as a responsive answer to the considerations that motivate nominalistic commitments in the first place? Indeed, it obliges us to ask the question: What do nominalists want? What are the rules of their game?

It cannot be that nominalism consists in insisting that all we do is refer to particulars using singular terms. Nominalists must allow that we also *say* things. Doing that is more than merely referring to things. Even in the simplest case, it is saying something *about* the particulars we refer to. It is classifying those particulars somehow. Classification involves some kind of repeatability on the part of the classifiers. Leo and Leona are both lions, and they are both tawny. Leo and Leona are classified together in that one can correctly say “...is a lion” and “...is tawny” of the two of them. Sellars thinks of explaining what we are saying when we say

---

<sup>152</sup> Sellars is happy to put his claim more baldly: “[A]bstract entities which are the subject of contemporary debate between platonist and anti-platonist philosophers—qualities, relations, classes, proposition, and the like—are linguistic entities.” [AE I, p. 163]. In the next section, I’ll give reasons why we should resist this formulation.

that as a modern version of the classical “problem of the one and the many.” The beginning of wisdom in the area, for Sellars, is to distinguish that problem from the problem of universals: the problem of saying what properties are. His analysis

requires us to hold that not all *ones* over and against *manys* are universals (i.e. qualities, relations, sorts, kinds, or classes), and consequently to conclude that the problem of “the one and the many” is in fact broader than the problem of universals...<sup>153</sup>

That is, Sellars will distinguish a narrower class of abstract entities, what he calls “universals”, from a broader class. He offers a deflationary metalinguistic nominalist analysis only of the narrower class. I will call this the strategy of distinguishing two grades of involvement in abstraction.

Following Carnap, Sellars is an ontological nominalist because he is a semantic nominalist. (And I will argue further along that that semantic deflationism is rooted in conceptual dependencies at the level of pragmatics—that is, in deep features of the *use* of the expressions addressed.) Here is a crude initial statement of the line of thought. Nominalism, as its name suggests, begins with views about *names*—or more broadly, singular terms. What there is can be named. (That is the connection between ontology and semantics, for nominalists of the sort under discussion.) What appear to be property-names or kind-names are not genuine names. So there are no such things. Sellars takes it, though, that common nouns, sortal expressions, are part of the apparatus of naming. For singular terms require criteria of identity and individuation that are supplied by covering sortals. The sortals also supply basic criteria and consequences of application for those singular terms (distinguishing them from mere labels).<sup>154</sup> Those sortals are, accordingly, a kind of “one in many” with respect to the objects that are referents of singular terms they govern. By contrast to

---

<sup>153</sup> *AE I*, p. 166.

<sup>154</sup> Sellars discusses this distinction in *CDCM*:  
 ...although describing and explaining (predicting, retrodicting, understanding) are *distinguishable*, they are also, in an important sense, *inseparable*. It is only because the expressions in terms of which we describe objects... locate these objects in a space of implications, that they describe at all, rather than merely label. [§108]  
 I talk about it in Chapter Eight of *Reason in Philosophy* [Harvard University Press, 2009].

the narrower class of universals, this, Sellars thinks, is a kind of one in many that the nominalist cannot and should not do without. He says:

[T]o refer to such a *one* we need a singular term other than the singular terms by which we refer to individual pawns, and yet which does not refer to a universal of which they are instances.<sup>155</sup>

If sense can be made of this kind of unity in diversity, then the way is open to understanding linguistic expression-types on this model, rather than on the model of universals and their instances or exemplifications. Doing so provides a way of responding to the second large objection to metalinguistic nominalism.

For a paradigm of a “one against a many” that is *not* a universal, not an abstract entity in the narrower, objectionable sense, he offers *distributive singular terms* (DSTs), such as “the lion” or “the pawn”. We can use them to say such things as “The lion is tawny,” and “The pawn cannot move backwards.” These can be understood as paraphrases of “Lions are tawny,” and “Pawns cannot move backwards.” These latter are things one understands as part of understanding how to use the common nouns, which is already part of understanding the use of singular terms such as ‘Leo’. Here is the strategy:

If, therefore, we can understand the relation of *the lion* (one) to *lions* (many) without construing *the lion* as a universal of which lions are instances; and if the looked-for singular term pertaining to pawns can be construed by analogy with “the lion”—indeed, as “the pawn”—then we would be in a position to understand how *the pawn* could be a one as against a many, without being a universal of which pawns are instances. This in turn would enable a distinction between a generic sense of “abstract entity” in which the lion and the pawn as well as triangularity (construed as the •triangular• ) and that two plus two equals four (construed as the •two plus two equals four• ) would be abstract entities as being ones over and against manys and a narrower sense of abstract entity in which qualities, relations, sorts, classes, propositions and the like are abstract entities,

---

<sup>155</sup> AE I, p. 166.

but of these only a proper subset, universals but not propositions, for example, would be *ones* as over and against *instances* or *members*. This subset would include the kind *lion* and the class of pawns, which must not be confused with *the lion* and *the pawn* as construed above.<sup>156</sup>

The contrast between two levels of involvement in abstraction is then the contrast between two sorts of nominalizations of common nouns such as “lion”, “pawn”, and •triangular•. Nominalizing common nouns (deriving singular terms from them) in the form of DSTs such as “the lion” is perspicuous and nominalistically unobjectionable, while nominalizing them to form kind-terms, such as “lionhood” is not. I want to propose that one lesson that can be drawn from Sellars’s is that **we can understand nominalism in terms of differential attitudes towards different kinds of nominalization**. But we will have to work our way up to this point.

The capacity to use distributive singular terms can be algorithmically elaborated from the capacity to use the common nouns they are derived from, via the schema

The K is F  $\equiv$  Ks are F.

The right-hand side of this equivalence is not a conventional quantification. In the case of natural kind-terms, like “lion”, it is something like essential properties that matter. The claim about Ks can be thought of as modified by something like Aristotle’s “generally, or for the most part” operator. (The existence of a non-tawny lion would not falsify “The lion is tawny.”)<sup>157</sup> The case we really care about, DSTs formed from common nouns formed by dot-quoting expressions, has special features, however. Sellars introduces them by analogy to “the pawn”, rather than “the lion.” The features that determine the truth of statements of the form F(the pawn) (“The pawn cannot castle,”), he says, are *prescriptive* rather than *descriptive* features of pawns. He means that it is the normative features that define the role something must play in a game to be a pawn—what features of its behavior are obligatory or permissible for pawns—that determine the truth-value of statements in which the DST occurs essentially. Besides those properties, each pawn will have matter-of-factual properties, such as being carved of wood, or being less than one inch tall, which are contingent features of some realizers, some items that play the role of pawn. Those do not

---

<sup>156</sup> AE I, p. 167.

<sup>157</sup> What I say here should be understood as only a crude gesture at a complex and important topic. For a more nuanced discussion, see Part One of Michael Thompson’s pathbreaking *Life and Action* [Harvard University Press, 2008].

support statements using the DST “the pawn.” In this respect, “the pawn” is like “the •triangular•”. It is norms governing the use of •triangular•s that determine what is true of the DST, too—even though “the pawn”, unlike “the •triangular•” is not metalinguistic.

The equivalence schema shows that DSTs are just a special way of referring to Ks: to lions or to pawns. Not to one single K, but to all of them, distributively. That the reference is distributive means that it is not to the *group* of Ks, but, as it were, to Ks *as* Ks.<sup>158</sup> We can contrast this special mode of distributive reference with another bit of technical machinery that has been used by another kind of nominalist (Goodmanian nominalists) to do some of the same work Sellars wants DSTs to do: mereology. Mereological sums, too, are “ones in many.” And they are different from universals. The part-whole relation they stand in to their mereological parts is not that of kind or property to *instance*. The difference is that mereological sums are a special kind of thing, over and above their parts. Singular terms referring to such sums are not special ways of referring to the parts, as DSTs are for particulars to which the common nouns from which they are formed apply. In this respect, mereological nominalism is *less nominalistic* than Sellarsian metalinguistic nominalism. For DSTs are not construed as singular terms referring to a different kind of entity from ordinary particulars. The mode of reference is different, specifically, distributive. But what is referred to is just what common nouns apply to. And that is the same particulars that singular terms refer to. There is no appeal to things of other ontological categories besides particulars. By contrast, mereological sums are formed from their parts by abstraction, as sets are. The difference between mereological sums and sets lies in the equivalence relation that is the abstractor, not in their abstractness.<sup>159</sup> Sellarsian nominalism must regard mereological sums, no less than sets, as ultimately metalinguistic in character.

The case Sellars really cares about, of course, is where the common nouns from which DSTs are formed are themselves the result of dot-quoting expressions of some type. An instance of the DST equivalence is:

The •triangular• is a predicate  $\equiv$  •triangular•s are predicates.

---

<sup>158</sup> Sellars says remarkably little about just how he thinks plural statements such as “Lions are tawny,” in terms of which statements formed using DSTs, such as “The lion is tawny,” are to be understood. He might have only a slippery grip on the point that what is true of “the mayfly” can be quite different from what is true of most mayflies. Michael Thompson offers a sophisticated discussion of this point in *Life and Action* [Harvard University Press, 2008]. Ruth Millikan’s notion of Proper Function underwrites quite a different analysis of the same phenomenon.

<sup>159</sup> Cf. the discussion in Chapter Six.



And, given Sellars’s analysis of property-names, we can extend this to:

The ●triangular● is a predicate ≡  
 ●triangular●s are predicates ≡  
 triangularity is a property.

Unlike “the lion” and “the pawn”, “the ●triangular●” is a *metalinguistic* DST. It refers, distributively, to expression-types (in a variety of actual and possible languages). That is why this Sellarsian analysis is, like Carnap’s less sophisticated account, a *metalinguistic* nominalism about what is expressed by property-names as a subset of ontological category vocabulary. Triangularity-talk is understood to be a misleading (because not explicitly metalinguistic) way of talking about the ●triangular●, that is, ●triangular●s, that is, expression-types that stand to “triangular” in some suitable (not for these purposes fully specified) relation of functional equivalence.<sup>160</sup> The equivalence relation is not, however, being appealed to as an abstractor that yields a singular term referring to an abstract object (perhaps identified with the equivalence class) that stands to the things it is abstracted from in a relation of exemplification. This is the difference between talking about the lion, or just lions—which is a way of referring to lions—as opposed to lionhood.

That is the difference between two kinds of ones-in-many, which is the basis of Sellars’s response to the objection that metalinguistic nominalism about properties and kinds must just trade nonlinguistic universals for linguistic ones. The strategy of distinguishing two grades of involvement in abstraction does trade nonlinguistic universals (lionhood, triangularity) for linguistic ones-in-many (the ●lion●, the ●triangular●), but not for linguistic *universals*. The explanatory progress being made corresponds to crossing the line between two sorts of unity in diversity. Universals (properties, kinds) are eschewed entirely.

---

<sup>160</sup> I have suppressed niceties concerning Sellars’s distinction, in *AE*, between “triangular” and ★triangular★ (the first being a quote-name of a word type, the second a quote-name of a sign-design type. Expressions formed by dot-quoting are officially common nouns applying to the latter, not the former.

#### IV. Nominalism and Nominalization, Functions and Objects

I said above that a metalinguistic nominalism that relies so heavily on this distinction between different kinds of repeatables—abstract entities in a strict or narrow sense where singular terms and covering common nouns are introduced by abstraction using equivalence relations on their instances, and divided (distributive) modes of reference to particulars—raises questions about the motivation for nominalism of this sort. Nominalism can be thought of as a hygienic recommendation regarding the conditions under which it is appropriate to introduce names—or, more generally, singular terms. More particularly, I think it is useful to think of nominalism as a policy concerning *nominalization*: the introduction of new singular terms (and common nouns or sortal expressions governing them) by grammatically transforming other expressions.

Sellars is concerned to distinguish two ways of nominalizing common nouns. “Lion” can be nominalized by abstraction, to form the property-name “lionhood.” Or it can be nominalized by forming the distributive singular term “the lion,” which we can understand in terms of the *plural* “lions.” The basic claim of this sort of nominalism is that nominalizations of the former sort are unspicuous and misleading, requiring metalinguistic analysis in terms of operators that form common nouns applying to expression-types by dot-quoting expressions illustrating those types, and operators that form DSTs from those dot-quoted expressions. (Abstractive nominalizations are “quasi-syntactic,” that is, material mode versions of statements perspicuously framed in the formal mode, as Carnap describes them in *The Logical Syntax of Language*. Sellars’s corresponding term is “covertly metalinguistic.”) Nominalizations of the latter sort are all right as they stand. Adjectives such as “...is triangular” and “...is red” take only nominalizations of the misleading abstractive sort: “triangularity” and “redness.” Nominalism is a set of

scruples about nominalization—a division of nominalization strategies into acceptable and unacceptable, or at least perspicuous and unperspicuous.

Although my focus here has been on predicate-nominalizations and properties, Sellars also thinks that declarative sentences have only nominalizations of the narrow sort, which purport to name abstract entities in the form of propositions. He proposes that these be analyzed metalinguistically, by equivalences of the form:

“That snow is white is a proposition.” ≡

“The •Snow is white• is a sentence.” ≡

“•Snow is white•s are sentences.”

So an extensional characterization of the split between nominalizations that unperspicuously invoke abstracta in the narrow sense (which are to be analyzed metalinguistically, using dot-quotes and DSTs), and nominalizations that invoke ones-in-many that are not covertly metalinguistic is this: kind-terms (sortals, common nouns) can go either way, depending on what sort of nominalization is at issue. Predicates (adjectives) and declarative sentences only take nominalizations that seem to refer to abstract entities in the narrow sense, and are to be understood by deflationary metalinguistic paraphrases. The only categories of expression-types that admit of nominalizations that are not to be construed as covertly metalinguistic are singular terms themselves (which are, as it were, their own nominalizations) and common nouns.<sup>161</sup> What is the motivation for this way of distinguishing the two grades of involvement in unperspicuous abstraction?

I said above that for the metalinguistic nominalist, the reason common nouns take nominalizations that are not covertly metalinguistic (such as “the lion” and “lions”), is that they are already involved in the mechanism of singular reference to particulars—that is, broadly speaking, in naming. They also take unperspicuous, covertly metalinguistic nominalizations, purporting to name abstract entities in the narrow, objectionable, sense, (such as “lionhood”) because besides incorporating criteria of identity and individuation (permitting plurals and so distributive reference) they are like predicates in incorporating criteria and consequences of

---

<sup>161</sup> For a possible qualification, see the remarks about gerunds (present participles) at the end of Section VI.

application. This means common nouns come with associated predicate-adjectives (“...is a lion”), which admit nominalizations purportedly naming abstract entities in the narrow sense the metalinguistic nominalist is concerned to deflate. But the reason common nouns also take nonmetalinguistic nominalizations must be that they can be construed as mechanisms of reference to particulars, albeit in the distinctive mode of plural, divided, or distributive reference, not just that there can be no singular term reference in the absence of individuating sortals. For it is equally true that there can be no singular term reference (“naming”) in the absence of assertion of declarative sentences (“saying”) or (therefore) predicating. Yet nominalizations of expression-types of those grammatical categories admit only ontologically unobscure nominalizations.

At the end of “Abstract Entities” Sellars offers a further characterization of the difference between abstract entities in the narrow sense, invoked by unobscure nominalizations to be nominalistically paraphrased metalinguistically, and in the wider sense. It corresponds, he says, to the distinction between abstract entities which are not *objects*, but *functions*.<sup>162</sup> He explicitly mentions Frege in this connection (while denying that there is anything paradoxical about reference to functions). Kind-terms (which have both criteria of application and criteria of individuation and identity) admit both readings, while predicate adjectives (which have only criteria of application) initially support only the functional reading. (They do admit of nominalizations that refer to objects, as we see below, but these are doubly unobscure and covertly doubly metalinguistic.)

The possibility that the word “kind” might have these two senses throws light on Russell’s erstwhile distinction between classes as ones and classes as manys. Or, with an eye to Frege, we can say that in contexts such as [“The •the lion• is a DST,” which reduces to “•the lion•s are DSTs”] kinds are *distributive objects*, whereas in [“The •lion• is a common noun,” which in turn reduces to “•lion•s are common nouns” (Sellars’s paraphrase of “Lionhood is a kind,”)]-like contexts they are concepts or functions.<sup>163</sup>

---

<sup>162</sup> *AE* VII, pp. 188-189.

<sup>163</sup> *AE* V, p. 186.

Again, he offers as examples:

Triangularity is a quality and not a (distributive) individual (i.e., The •triangular• is a predicate and not a DST).

Triangularity is a (distributive) individual and not a quality (i.e., The •the •triangular•• is a DST and not a predicate).<sup>164</sup>

Triangularity as a quality is a paradigm of a *function*, while *triangularity* as a distributive individual is a corresponding object. (Sellars marks the difference by using italics in the latter case.<sup>165</sup>) This sort of derivative nominalization corresponds to meta-metalinguistic DSTs.

While it is not immediately clear what Sellars means by saying that some of these nominalizations refer to functions rather than objects (and the invocation of Frege's views from "Concept and Object" and "Function and Concept"<sup>166</sup> threatens to explain *obscurum per obscurius*), it does seem that he is lining up abstract entities in the *narrow* sense with *functions*. Nominalizations that invoke functions are the unobscure ones (cf. "classes as ones"), by contrast to nominalizations that invoke objects, albeit distributively (cf. "classes as manys").

## V. Saying, Naming, and Predicating

---

<sup>164</sup> AE VII, p. 189.

<sup>165</sup> AE IV, p. 183-184.

<sup>166</sup> In Peter Geach and Max Black (trans.) *Translations from the Philosophical Writings of Gottlob Frege* [Blackwell, 1966], originally published in 1952.

I think Sellars explains his reasons for drawing where he does the line between nominalizations of the two kinds—straightforward and covertly metalinguistic—and for the appeal to a distinction between objects and functions, in the third of the trio of essays I have been considering, “Naming and Saying.” The proximal topic of this essay is the contrast between two different approaches to universals: that of Gustav Bergmann (of the Vienna Circle) and one Sellars associates with Wittgenstein’s *Tractatus*.<sup>167</sup> Of particular interest is that accounts of both sorts end by appealing to something *ineffable*—though the ineffability arises at characteristically different places in the two. Though himself coming down firmly on the Tractarian side of the dispute, as he understands it, Sellars diagnoses the objectionable ineffability as having a common aetiology in the two cases—as being rooted in the same failure of understanding.

In its crudest terms, the Bergmann-*Tractatus* debate is about how many ontological categories of things there are in the world, and how we should understand their relations. For Bergmann, there are two kinds of things, particulars and universals, and just one relation, *exemplification* of a universal by particulars, that they can stand in.<sup>168</sup> Saying that two particulars stand in some relation, for instance that Ludwig is subtler than Gustav, is producing names of the two kinds (names of particulars and names of a universal) in a way that conventionally conveys that they stand in the relation of exemplification. The disappointing addendum is that that relation is ineffable. Naming (nominalizing) it, for instance, ‘exemplification’, is at best of heuristic and not analytic value, since the relation between it and the particulars and universal it relates (e.g. Ludwig, Gustav, and the relation of being subtler than) would itself have to be understood as...exemplification. And then we are off to the races on a Bradleyan regress.

---

<sup>167</sup> There are many fine things in this essay that I shall not discuss. Two subtleties worthy of at least passing mention are: i) Sellars’s sensitive and judicious treatment of vexed interpretive question of exactly what stand the *Tractatus* takes on the intelligibility of multiple distinct *monadic* facts (since facts are “arrangements” of objects); and ii) the distinction between color and shape predicates in this context: “green” has both adjectival and substantival uses, which invites confusion (it can serve as its own adjective-nominalization—“Green is a color”—though it also takes “greenness”), whereas “triangular” nominalizes *only* as “triangularity”.

<sup>168</sup> Sellars: “[F]or Bergmann there is...only *one* relation, i.e. exemplification, and what are ordinarily said to be relations, for example *below*, would occur in the world as *relata*.” [NS 109]

By contrast, according to the Tractarian view Sellars considers, there is only one kind of thing in the world: particulars. They stand in a variety of relations. Saying that two particulars stand in some relation, for instance that Ludwig is subtler than Gustav, is arranging names of the particulars in a way that conventionally conveys the fact that the particulars stand in that relation. The disappointing addendum is that the relation (picturing) between statement (the fact that the names are arranged as they are in the saying) and the fact (that the particulars stand in the relation) is ineffable. It is not itself a fact that can be stated, as a relation obtaining between a names-fact and a particulars-fact, but only something that can be shown. Here what threatens is not so much a regress as circularity: the explicit statement of the semantic picturing relation between statements and facts could be understood only by someone who already implicitly grasps the relation between statements and facts, and so could not substitute for or ground such a grasp.

Here is Sellars's summary:

To keep matters straight, it will be useful to introduce the term 'nexus' in such a way that to say of something that it is a nexus is to say that it is perspicuously represented in discourse by a configuration of expressions rather than by a separate expression. If we do this, we can contrast Bergmann and Wittgenstein as follows:

*Wittgenstein:* There are many nexus in the world. Simple relations of matter of fact are *nexus*. All objects or individuals which form a nexus are particulars, i.e. individuals of type 0. There is no relation or nexus of exemplification in the world.

*Bergmann:* There is only one nexus, exemplification. Every atomic state of affairs contains at least one...individual which is not a particular.

If one so uses the term 'ineffable' that to eff something is to signify it by using a name, then Wittgenstein's view would be that what are ordinarily called relations are ineffable, for they are all nexus and are expressed (whether perspicuously or not) by configurations of names. For Bergmann, on the other hand, what are ordinarily called relations are effed; it is exemplification which is ineffable.<sup>169</sup>

---

<sup>169</sup> [NS 109].

Notice that Sellars here expresses the nominalism being opposed to Bergmannian ontological profligacy as a restriction on what can strictly be *named* (hence how nominalizations are to be understood: where straightforwardly and where in terms of metalinguistic paraphrase). An assumption taken to be common to all concerned is that what can be named and what is “in the world” coincide, and that anything else is strictly “ineffable”. One might rather tie ineffability to what cannot be *said* (explicitly) but at most only shown or otherwise conveyed (implicitly). I’ll return to this question.

Sellars sensibly takes the invocation of something ineffable as a symptom of analytic and explanatory failure. His diagnosis (repeated with emphasis in the concluding sections of both *NS* and *AE*) is that the surplus beyond what is named when we say something, what shows up on these mistaken accounts as ineffable, is not a *thing* but a *doing*.

Thus the “relation” of exemplification which for Platonists binds the realm of becoming to the realm of being, and which for more moderate realists binds the “real” order to the “logical” or “conceptual” order, is an offshoot of the “relation” of truth, which analysis shows to be no relation at all, but a sign of something to be **done**. [*AE* 203]

The supposedly ineffable alternatives, exemplification (Bergmannian platonism) and the relation between statements and facts (Tractarian nominalism) are both manifestations of what is invoked by truth-talk. And that, Sellars thinks, is best understood not in terms of a word-world relation but in terms of the propriety of a metalinguistic *inference*.

What, then, does it mean to say

That green a is a fact

Clearly this is equivalent to saying

That green a is true

...

This, however, is not the most perspicuous way to represent matters, for while the equivalence obtains, indeed necessarily obtains, its truth depends on the principle of inference—and this is the crux—

From ‘that green a is true’ (in our language) to infer ‘green a’ (in our language).

And it is by virtue of the fact that we *draw* such inferences that meaning and truth talk gets its connection with the world. In this sense, the connection is *done* rather than *talked about*.



Viewed from this perspective, Wittgenstein's later conception of a language as a form of life is already foreshadowed by the ineffability thesis of the *Tractatus*. But to see this is to see that no ineffability is involved. For while to infer is neither to refer to that which can be referred to, nor to assert that which can be asserted, this does not mean that it is to fail to effect something which is, therefore, ineffable." [NS 125]

A number of moves are being made here. First, the "two ineffables", exemplification and the relation between statements and facts, are both being traced back to what is expressed by statements using 'true'. "a *exemplifies* green" is a way of *stating* the *fact* that a is green. (Stating is the paradigmatic kind of saying.) Second, "A fact is a thought that is true" (Frege, in *The Thought*).<sup>170</sup> (Keep in mind the "notorious 'ing'/'ed' ambiguity here: he does *not* mean 'thought' in the sense of a thinking, an act, but in the sense of what is thought—or better, thinkable—a content.) Third, talk about truth is (as Frege also recognized), misleading talk about what one is *doing* in saying something in the sense of making a statement: the use of 'true' is to be understood in terms of the platitude that asserting is taking-true. Fourth, the way 'true' expresses what one is doing in asserting is also expressed in the propriety of the disquotational *inferences* codified in Tarskian T-sentences.

All of these moves are contentious. I am not concerned to defend them here. I am concerned to understand the original motivation and general rationale for connecting nominalizations the Sellarsian nominalist wants to treat as not referring to *things*, such as "triangularity", with discursive *doings*. For this, I want to suggest, is what becomes of the otherwise puzzling distinction, evidently intended to be coextensional, which we worried about at the end of the previous section, between referring to *objects* and invoking *functions*. As we might break things down, in the first step, functions are what articulate functional roles. In the second step, functions, as Sellars is thinking of them, are things only in the sense of things *done*: doables. Nominalization of functions is what Sellars's nominalism invites us to forbid in perspicuous languages, and to give a deflationary treatment of the functioning of, in unperspicuous ones.

I think we can begin to understand the idea behind this line of thought if we look at the activities of that give "Naming and Saying" its title, and how the relations between them are

---

<sup>170</sup> "The crucial ineffability in the *Tractatus* concerns the relation between statements and facts. Is there such a relation? And is it ineffable? The answer seems to me to be the following. There is a meaning relation between statements and *facts*, but both terms are in the linguistic order." [NS 124]

thought to be made more perspicuous by the third technical innovation (besides dot-quotes and DSTs) that Sellars uses to articulate his nominalism. This construction, introduced in that essay, is the language-form he calls “Jumblese.”<sup>171</sup> We can sum up the line of thought in *NS* that I have been considering in the slogan: Appeal to an ineffable semantic relation is a sign that one is trying to do in one’s *semantic* theory what can only be done in the *pragmatic* theory, the theory of the *use* of the language. *Saying*, putting something forward *as* true, asserting—the central and paradigmatic use of declarative sentences—is a *doing*, not a semantic relation. So is *naming*, in the sense of referring (using an already established term, rather than naming in the sense of introducing such a term). Referring is the central and paradigmatic use of singular terms.

If the first lesson Sellars wants us to learn is that the result of trying to explain what one is *doing* in *saying* something (a pragmatic matter), in terms of the semantic relation between a name and what is named, is an appeal to an ultimately magical, ineffable version of that relation, then the second, nominalist, lesson is that even within the realm of semantics, the name/named model cannot be used to understand the use of predicates or sentences. In particular, predication, in the sense of the act of predicating (classifying something nameable) is a *derivative* speech act. It does not belong at the same level of analysis as the more fundamental acts of saying and naming. Predicating something (universal) of something (particular) is just saying something *about* something. It is to be understood in terms of the relation between a kind of *doing*, asserting, which in the base case essentially involves the use of singular terms, and the semantic relation of referring, which holds between a name (singular term) and what is named (referred to).<sup>172</sup>

---

<sup>171</sup> The name comes from Edward Lear’s nonsense poem “The Jumbles”, Sellars tells us, because “Far and few, far and few, are the lands where the Jumbles live.” (He does not mention that “Their heads are green, and their hands are blue...”, though his topic is the significance of just such predications. Greenness and blueness are not mentioned on the inventory of things they took with them when they “went to sea in a Sieve.”)

<sup>172</sup> Though he does not say so, I expect that Sellars learned from Kant the lesson that one cannot, as the pre-Kantian tradition tried to do, understand *saying* in terms of *predicating*. I explain how I take Kant to have learned this lesson, and the central role it plays in his thought, in Chapter One of *Reason in Philosophy*.

It is because the speech act of predicating is a derivative one that predicative expressions play a subordinate role to singular terms and sentences.

[T]he classical problem of universals rests in large part on the fact that, in such languages as English and German expressions referring to universals are constructed on an illustrating principle which highlights **a design which actually plays a subordinate role**, and consequently tempts us to cut up such sentences as

Triangular (*a*)

into two parts, one of which has to do with the universal rather than the particular, the other with the particular rather than the universal, and tempts us, therefore, to construe the statement as asserting a dyadic relation (“exemplification”) to obtain between the particular and the universal.<sup>173</sup>

Jumblese is designed to make *syntactically* vivid the derivative *pragmatic* role of predication, which in turn underlies the deflationary, nominalist metalinguistic *semantic* analysis Sellars is recommending for *nominalizations* of predicative expressions, such as “triangularity.” Jumblese has no predicative expressions. Its sentences consist entirely of names (singular terms). The names specify what one is talking *about* (referring to). What one is *saying* about what one is talking about is expressed by *styles of arrangement* of those names. So, in one version the claim that Wilfrid is subtler than Gustav might be expressed by juxtaposing them and writing the first name in larger type than the second: Wilfrid Gustav. That Gustav was Austrian might be expressed by writing his name in a distinctive font: *Gustav*. Jumblese, we might want to say, overtly marks only naming and saying: what one is referring to, by the singular terms used, and what one is asserting about it, by the style in which the terms are written (including the relations between the singular terms). Predication is only implicit in what one is doing in saying something about something.

A consequence of the absence of overt predicate-expressions is that there is nothing to nominalize into an analog of “triangularity.” There is nothing to which to apply the “illustrating principle” that forms •triangular•, which could tempt one to introduce the new common noun “property”, enabling one to say “Triangularity is a

---

<sup>173</sup> [AE 201].

property,” that is, •triangular•s are predicates (the •triangular• is an adjective). Of course, we *could* introduce nominalizations of predicate-adjectives even into (a dialect of) Jumblese, perhaps by using names of the styles the level-0 names are written in. Since it is the *fact that* “*Gustav*” is written in the Script-MT-Bold font that *says that* Gustav is Austrian, we could say that •...is Austrian•s are predicates (that is that being Austrian is a property) by saying that Script-MT-Bold is a predicate-indicating font—or, in a Jumblese metalanguage, by asserting “**Script-MT-Bold**” (where writing the font-name in the Berlin Sans FB font indicates that it is the nominalization of a predicate).<sup>174</sup> But while Jumblese *permits* such nominalizations, it does not *encourage* them. And it does not even *permit* the formation of those nominalizations according to an *illustrating* principle, which is what makes ontological-category talk such as “Triangularity is a property” *covertly* metalinguistic (Carnap’s “quasi-syntactic”): a formal-mode statement masquerading in material mode. “**Script-MT-Bold**” is *overtly* metalinguistic, consisting, as it does, of a name of a style of writing, here, a font (itself, of course, written in a particular style).

## VI. From Semantic to Pragmatic Metalanguages: Assessing Metalinguistic Nominalism

In the earliest of the three essays I have been discussing, Sellars identifies two major objections to Carnap’s metalinguistic nominalism about ontological category vocabulary, principally predicate-nominalizations (such as “triangularity”) and their associated common nouns (such as “property”). First, statements such as “Triangularity is a property,” do not mention any linguistic expressions, and so are not metalinguistic in the classical sense. Unlike Carnap’s proposed paraphrase, “‘Triangular’ is a predicate,” they would be true even if no-one had ever spoken English, and do not change their reference or become unintelligible to monolinguals if translated into German. Second, it seems such an approach just trades non-linguistic universals, such as “being triangular” for linguistic ones, such as “being a predicate.”

---

<sup>174</sup> In Section VIII of *AE*, Sellars considers how bound variables might work in Jumblese. (But do his readers care? The result of this expository choice is an *extremely* anticlimactic ending—one could not say conclusion—to the already long and technical essay.) Elsewhere in the same piece, he indulges himself in speculations about Jumblese metalanguages (*inter alia*, for Jumblese), and about the adventure that would consist in translating Bradley’s *Appearance and Reality* into Jumblese. Oddly, he says nothing about the *spoken* version of Jumblese—the version in which, we are authoritatively informed, the Jumbles said “How wise we are! Though the sky be dark and the voyage be long....” One version of spoken Jumblese would be tonal: melodic. The effect would be reminiscent of Gregorian chants. A written Jumblese *pragmatic* metavocabulary for such spoken Jumblese would resemble musical notation (and *its* use, a Glasperlenspiel).

Sellars's response to the first objection is that it turns on too narrow and undifferentiated a conception of the metalinguistic. He offers a more capacious and nuanced one, reformulating Carnap's paraphrase using dot-quotation to form common nouns that functionally classify expression-types using the "illustrating sing-design principle." He responds to the second by conceding that classification under repeatables is not to be explained away, but insisting that we should distinguish the broader "problem of the one and the many" from the narrower "problem of universals." The formation of plurals from common nouns (including those formed by dot-quotation of illustrating expressions: "•triangular•s are predicate-adjectives") and their nominalization by forming distributive singular terms instead of kind-names ("the •triangular•" rather than "•triangular•ness"—in the non-metalinguistic case, "the lion" rather than "lionhood") allows the metalinguistic nominalist to endorse a version of Carnap's paraphrase without commitment to linguistic (or any) universals in the narrow, objectionable sense.

I think these responses are wholly successful in producing a development of Carnap's idea that is immune to the objections that prompted them. The second move, however, prompts the question of why we should resist reifying universals in the form of properties and kinds. Why should we insist on metalinguistic paraphrases of claims made using *these* nominalizations, and hence reject a straightforward referential semantics for these singular terms, which understands them as referring to abstract entities? Sellars's argument, as presented in "Naming and Saying," turns on the second-class ("derivative", "subordinate") character of *predicating* (and, more generally, classifying), relative to saying and naming. That is, the basis for metalinguistic nominalism about property and kind nominalizations in *semantics* is to be found in considerations proper to *pragmatics*: considerations concerning what we are *doing* when we use various expressions. I think we can and should resist this move.

Sketched with a very broad brush, I think the argument goes like this. Predicate-adjectives have a very different function and use than do singular terms. *Hence*, it is misleading to understand singular terms formed by nominalizing them as referring to a special kind of *thing*:

abstract entities.<sup>175</sup> I don't think this is a good inference. It is true both that predicating is not naming, but must be understood in terms of the relations between naming and saying, and that one can only understand singular terms formed by nominalizing predicates in terms of the use of the underlying predicates. On this latter point, Sellars argues in effect that the capacity to use ontological category talk—predicate- and kind-nominalizations, such as “triangularity” and “lionhood”, and the common nouns that govern their identity and individuation, such as “property”<sup>176</sup> and “kind”—is *pragmatically dependent* on the capacity to use the underlying predicate-adjectives and common nouns. In the terms I use in *Between Saying and Doing*, this is a PP-necessity claim.<sup>177</sup> Unless one has the capacity to use the nominalized terms, one cannot count as having the capacity to use their nominalizations. Further, his version of the Carnap metalinguistic paraphrase strategy shows us how the capacity to use predicate-adjectives (“...is triangular”) can be *algorithmically elaborated* into the capacity to use the nominalizations (“triangularity”).<sup>178</sup> This is a special kind of PP-sufficiency claim. I agree with all this, and think that showing how to algorithmically elaborate the ability to use adjectives into the ability to use nominalized property-talk is a significant achievement. Further, I agree that the pragmatic dependence (PP-necessity) claim suffices to show that Bergmann is wrong to think of the nominalization-talk as *conceptually prior* to the use of the predicate-adjectives and ground-level

---

<sup>175</sup> A corresponding argument goes through for common nouns, which are like predicate-adjectives in having classifying criteria of application, even though they are unlike predicate-adjectives in also having criteria of identity and individuation for the singular terms associated with them. Also, Sellars wants to adopt the same sort of metalinguistic paraphrase strategy for nominalizations of sentences (“that snow is white,” together with the corresponding common nouns such as “proposition”). Again, the avowed motivation for this is that what one is *doing* in *saying* something is different from *referring* (though referring to particulars is in the base case included as one aspect of saying). Nonetheless, for simplicity, in this summary, I focus on the predicate-adjectives and their nominalizations.

<sup>176</sup> And, though he doesn't say so, others such as “trope”, understood as something like “unrepeatable instance of a property.”

<sup>177</sup> Oxford University Press, 2008. Hereafter “*BSD*”.

<sup>178</sup> Sellars suggests that the fact that some kind-terms mark *functions* rather than *objects* (discussed in Section IV above) means that thinking of them as naming universals is committing something like the naturalistic fallacy. In this respect, he seems to be putting abstract-entity-talk in a box with normative vocabulary. Normative vocabulary, like modal vocabulary, he takes to play the expressive role, not of describing something (“in the world in the narrow sense”), but of explicating the *framework* within which alone describing is possible. (I discuss this Kantian move in Chapter Five.) These vocabularies are what in *BSD* I call “universally LX”: elaborated from and explicative of every autonomous vocabulary. I have just been claiming that the use of ontological-category vocabulary (such as “property” and “proposition”—the common nouns that govern singular terms purporting to pick out abstract objects such as universals like triangularity) can indeed be elaborated from the use of ordinary predicates and declarative sentences. One very important question that I do not address in this essay is whether (for Sellars, and in fact) such vocabulary is *also* explicative of essential features of the framework within which ordinary empirical descriptive vocabulary functions, and if so, of *which* features.

common nouns. Bergmann is right that there is a *semantic* equivalence between saying that a exemplifies triangularity and saying that a is triangular. However, there is an underlying *pragmatic asymmetry*. One *could* learn how to use “...is triangular” (•triangular•s) first, and only then, and elaborated solely on that basis, learn how to use “...exemplifies triangularity”, and the property-talk that goes with it (as the common noun to this nominalization-by-abstraction). One could *not* learn it the other way around. In this sense, property-exemplification talk is not *pragmatically autonomous* from the use of predicate-adjectives, as Bergmann’s priority claim commits him to its being. This sort of *pragmatically mediated conceptual dependence* is the same sort of priority claim that Sellars makes for “is”-talk over “seems”-talk, in *EPM*.<sup>179</sup> So far, so good.

More particularly, Sellars’s claim is that what one is *doing* in saying that triangularity is a property is classifying •triangular•s as predicate-adjectives. That is a metalinguistic doing—of a distinctive kind, marked out by the use of the illustrating principle, to get a common noun, •triangular•, that applies to expression-types that stand to the displayed “triangular” in a parameterized functional-role equivalence relation. So it is fair to conclude that the *use* of ontological-categorial vocabulary involves a distinctive kind of metalinguistic expressive role. The question remains: what conclusions should one draw about the *semantics* of such expressions? Does playing that *pragmatic* metalinguistic expressive role *preclude* understanding the nominalizations (“triangularity”, “lionhood”—or “being a lion”) as *also* standing in referential (“naming”) relations to *objects*? I do not see that it does. The fact that “good” essentially, and not just accidentally, has as part of its expressive role the possibility of being used to *commend* does not mean that it does not *also* describe in the sense of attributing a property. A corresponding point goes through for *modal* vocabulary.<sup>180</sup> From that fact that what one is *doing* in *saying* that triangularity is a property is classifying •triangular•s as predicate-adjectives, it does not follow that that is what one is *saying*. It certainly does not follow that that is *all* one is saying. Sellars’s analysis leaves *room* for denying that “triangularity” refers to a property. It provides an alternative. But he has not shown that these are *exclusive* alternatives,

---

<sup>179</sup> See chapter One of *BSD*.

<sup>180</sup> As I argue in Chapter Five, “Modal Expressivism and Modal Realism, Together Again.”

that we must *choose between* them. The singular terms formed by nominalizing parts of speech other than singular terms are, we might agree, distinguished by having a metalinguistic expressive function. But that is not yet to say that they do not *also* refer to a distinctive kind of object: property-universals (and propositions, including the true ones: facts).

Traditional Tarskian metalanguages—the kind we normally think about in connection with “metalinguistic” claims—are *semantic* metalanguages. They contain the expressive resources to talk about aspects of discursive *content*. Accordingly, they let us discuss truth conditions, reference relations, inferential relations, and the like. Carnap also deploys *syntactic* metalanguages, that let us talk about syntax, grammar, and lexical items (though Carnap himself uses “syntax” in an idiosyncratically wide sense in *The Logical Syntax of Language*). *Pragmatic* metalanguages have the expressive resources to talk about the *use* of language and the proprieties that govern it, for instance the activities of asserting, inferring, referring, predicating, and so on.<sup>181</sup> If I am right that the principle insight driving Sellars’s metalinguistic nominalism is the idea that what one is *doing* in deploying concepts such as triangularity, lionhood, property, and kind is functionally classifying expressions using metalinguistic vocabulary of a distinctive kind (nominalizations formed according to the “illustrating sign-design principle”), that is an insight properly expressible in a *pragmatic* metalanguage. The conclusion he wants to draw, however, concerns the *semantics* proper for that class of nominalizations and covering common nouns. The inferential relations between claims couched in pragmatic metalanguages and claims couched in semantic metalanguages are quite complex and little understood, however.<sup>182</sup> The inference Sellars is committed to here would go through only in the context of one or another set of auxiliary hypotheses, many of which would be implausible, or at least controversial, none of which does he discuss.

Sellars makes this slide unaware (to be sure, in the good company of expressivists addressing other sorts of vocabulary) because he doesn’t have available the distinction between semantic and

---

<sup>181</sup> Pragmatic metavocabularies are one of the topics discussed at length in *BSD*.

<sup>182</sup> *BSD* introduces the topic, and provides a wealth of examples of the sort of complex relations between meaning and use that can be discerned once we start to think systematically about their relations.



pragmatic metalanguages. According to that diagnosis, his argument is vulnerable because it relies on too crude and expressively impoverished a concept of the metalinguistic. This is an ironic situation, because I am accusing Sellars of making a mistake (or suffering from a disability) of a piece with the ones he discerns in the opponents he discusses in these essays. As we have seen, the first principal objection to Carnap's metaphysical nominalism that Sellars addresses he diagnoses as the result of appealing to insufficiently nuanced concepts of being metalinguistic. He responds by giving us more nuanced ones, which evade the objection. I am claiming that his notion of the metalinguistic is still too crude. Again, he diagnoses Bergmann and the *Tractatus* as running together *pragmatic* issues, of what one is *doing* in saying something or *predicating* something, with *semantic* issues. In particular, he claims that attempting to understand what one is *doing* in predicating or claiming by forcing it into the form of a semantic relation inevitably results in commitments to the ineffability of that relation. This is the same genus as the mistake I am claiming he is making: running together *pragmatic* issues, of what one is *doing* in saying something, with *semantic* issues of what is *said* thereby.

This line of thought suggests that there are a number of different strands of broadly nominalistic thought in play. One genus is what might be called “nominalization nominalisms.” These views make an invidious distinction between two classes of singular terms. *Genuine* singular terms are *referential*. They are to be understood semantically in terms of reference relations (the “name-bearer” relation), and successfully using them is referring to a referent. Genuine singular terms in this sense can fail to refer, but they, as we might say, *perspicuously purport* to refer to particulars. They are not *grammatically* precluded from being used to refer, and in any case are to be semantically assessed in terms of reference relations (or the lack thereof). By contrast (almost all) singular terms formed by nominalizing other parts of speech are grammatically misleading. These merely *ostensible* singular terms only grammatically, but *unperspicuously purport* to refer to particulars. On Sellars's development of Carnap's view, they are to be given *metalinguistic* readings. All singular terms have criteria of identity and individuation lodged in associated common nouns or sortals, which accordingly can also be divided into genuine and ostensible. This division generally corresponds to that between nouns that are not, and those that are, formed by nominalizing other parts of

speech. The exception is that *some* nominalizations of common nouns or sortal expressions are sometimes counted as genuine (for instance, by Sellars and Kotarbiński).<sup>183</sup>

In these terms, I want to distinguish semantic and pragmatic species of the genus of nominalization nominalisms. The first kind of nominalization nominalism addresses the semantic *content* of the two classes comprising genuine and merely ostensible singular terms (the latter consisting of transcategorial nominalizations). Only genuine singular terms are to be understood in terms of their referential relations to particulars. The latter kind of nominalization nominalism addresses the pragmatic *use* of the two classes of terms and associated common nouns. The pragmatic nominalization nominalist understands the *use* of transcategorial nominalizations in metalinguistic terms of classifying linguistic expression-types. By contrast, the use of genuine singular terms is to be understood exclusively as *referring*, which is one essential feature of *saying* anything *about* particulars. I have claimed that the step from pragmatic to semantic nominalization nominalism is not straightforward. For one might distinguish transcategorial nominalizations from other singular terms by seeing their use as involving metalinguistic classification without thereby concluding that they do not *also* stand in referential relations to a distinctive kind of abstract entity. They just have this *extra* expressive function that *ordinary* singular terms do not have. Perhaps there is an illuminating and important relation between playing that distinctive expressive role and picking out the kind of object they do.<sup>184</sup>

In any case, when we discover that some kind of linguistic expression plays a distinctive expressive role (one not played by paradigmatically referring singular terms, for instance), we would seem methodologically to have two choices. We can think about that new expressive role in an *exclusionary* or in a *permissive* way. The exclusionary reading claims that the expressive role that has been discovered must exhaust what is available to determine semantic content. The contrasting permissive reading allows that playing that expressive role might be compatible with

---

<sup>183</sup> It seems, for the same reason: otherwise the nominalization nominalist about the “problem of universals” has difficulty addressing the “problem of the one and the many.” Kotarbiński, T., *Gnosiology. The Scientific Approach to the Theory of Knowledge*, trans. by O. Wojtasiewicz. [Oxford: Pergamon Press, 1966]. This is a translation of Kotarbiński’s *Elementy teorii poznania, logiki formalnej i metodologii nauk*, Lwów: Ossolineum, 1929. Kotarbiński distinguishes between “genuine” and “non-genuine” names, and between semantic and ontological nominalisms.

<sup>184</sup> I have in mind determining the equivalence relation that is the abstractor.

also playing other expressive roles (for instance, referring), and so not ruling out the corresponding semantics still being applicable. The fact that expressivists who want to adopt the exclusionary reading should argue for adopting this stance rather than the permissive one (as should those who want to adopt the less common permissive stance), of course, is not limited to the case of expressive nominalists.

What I have called “nominalization nominalisms” concern the use and content of linguistic expressions. Nominalism is usually thought of as an *ontological* thesis, however. Sellars endorses such a view:

It is also argued that exemplification is a ‘quasi-semantic’ relation, and that it (and universals) are “in the world” only in that broad sense in which the ‘world’ includes linguistic norms and roles viewed (thus in translating) from the standpoint of a fellow participant.<sup>185</sup>

I take it that being “in the world in the narrow sense” means being in the *nondiscursive* world: the world as it was before there discursive beings, or a counterfactual world in which there never were discursive beings. If this is indeed the narrow sense of “in the world” that contrasts with the broad sense invoked in this passage, then it seems to me that there is a tension between this claim and the response to one version of the first objection to naïve Carnapian metalinguistic expressivism about transcategorical nominalizations. This objection is that it cannot be right to understand sentences such as “Triangularity is a property,” metalinguistically, because they would still have been true even if there had never been discursive beings. Sellars’s response commits him to the claim that “•triangular•s are predicates” would still have been true even if there never had been discursive beings. Perhaps there are ways to vindicate this claim without being committed to •triangular•s being “in the world” in the narrow sense, but it is hard to see how. I suppose that he thinks that •triangular•s *are* “in the world in the narrow sense,” but that that is compatible with his claim, since •triangular•s are not universals and are not exemplified by the expression-types they classify. (They are “ones in many”, but not universals.) The presumptive presence of •triangular•s “in the world in the narrow sense” suggests that some work will need to be done to clarify and entitle oneself to appeal to this “narrow sense.”

---

<sup>185</sup> [NS 103].

Be that as it may, what is “in the world in the narrow sense” is being taken to exclude universals because they are not, as we first might have thought, referred to by *genuine* singular terms, but only by *ostensible* ones. Nominalism in the ontological sense is the thesis that the world (“in the narrow sense”) consists exclusively of *nameables*: things that could be referred to by *genuine* singular terms. This connection between *semantic* nominalism, which distinguishes genuine from merely ostensible <sup>s</sup>names<sup>s</sup> (singular terms), and *ontological* nominalism, which restricts the real to what is nameable by genuine ones, is explicit in Kotarbiński. It seems to be Sellars’s picture as well.

Now I am not at all sure that ontological nominalism in this sense is in the end so much as intelligible. In Sellars’s version of semantic nominalization nominalism, among the transcategorial nominalizations that are analyzed metalinguistically, and which accordingly show up as not genuine singular terms, are *sentence* nominalizations, and their associated common nouns such as “proposition” and “fact”. (“That snow is white is a proposition,” is analyzed as “•Snow is white•s are declarative sentences.”) Although “Naming and Saying” defends a Tractarian view against Bergmann on *some* important points, Sellars parts company with the *Tractatus* in taking a reistic position according to which the world (narrowly conceived) is *not* everything that is the case, a world of facts, but is rather a world exclusively of particulars, nameables not stateables. As far as I can see, Sellars is envisaging a world in which the “ones-in-many” needed to make sense of an *articulated* world are such as could be referred to by common nouns (sortals). That is the alternative to universals he seems to be working with. But to avoid commitment to universals, it seems that the criteria of identity and individuation associated with the (already, as it were, nominalized) common nouns must either *do all* the work, or must somehow immunize the criteria (and consequences) of application from supporting or making intelligible the contribution of the universals that threaten when predicate adjectives, which *only* have circumstances (and consequences) of application, but not criteria of identity and individuation, are nominalized. I don’t pretend to know that this strategy cannot be made to work. But I also don’t see that Sellars has given us many of the tools that would need to be deployed to make it work. Perhaps more fundamentally, I don’t see that we have the makings of a story on the ontological or the semantic side of what corresponds on the pragmatic side to *saying* (claiming, believing) something. If the world is a collection of particulars—of course, collections are not “in the world in the narrow sense” either—what is one doing in *saying that* things are thus-and-so? How are we to understand *either* the “thus-and-so” *or* the “saying that”? I am buffaloed.

Here is a potentially more tractable puzzle. I have interpreted the semantic side of Sellars’s nominalism as what I have called a “nominalization nominalism,” which distinguishes two classes of singular terms, genuine and merely ostensible. The merely ostensible ones are to be read metalinguistically, in the broad, nuanced sense of

“metalinguistic” that applies to DSTs formed from dot-quoted expressions using the “illustrating sign-design principle.” More specifically, I have claimed that all transcategorial nominalizations count for Sellars as merely ostensible singular terms according to this classification, and so, according to the ontological side of his nominalism, do not correspond to anything “in the world in the narrow sense.” One kind of transcategorial nominalization, starting with a non-nominal part of speech and forming singular terms from it, is gerunds or present participles, such as “doing”, “making”, “breaking”, “swimming”, and “heating”. These constructions form common nouns and singular terms from verbs. If my account of how the motivation of “Naming and Saying” shapes the account of “Abstract Entities” is correct—if being a transcategorial nominalization is sufficient for not being a genuine singular term for Sellars—then all singular terms formed from verbs must be merely ostensible, and correspond to nothing in the world construed narrowly. Sellars never discusses this case. Would he offer a broadly metalinguistic account of these terms and common nouns? If so, how would it go? Does his nominalism allow that the world “in the narrow sense” can include particular swimings and heatings? These seem like particular events, rather than universals. A particular swimming falls under the common noun “swimming” as a particular dog falls under the one-in-many “...is a dog,” rather than by way of exemplification. And the processes of Sellars’s late ontology can be thought of just as extended events, and seem naturally to be picked out by gerunds and present participles. So it seems that either there is a tension in Sellars’s nominalism on this point, or I have characterized his nominalization nominalism too broadly. But if that is so, how should we determine which nominalizations of verbs and adjectives are alright, forming genuine singular terms and common nouns, and which are not? The considerations of “Naming and Saying” do not seem to give us adequate guidance here.

I want to close with the observation that, putting aside the slide I have accused Sellars of making from pragmatic to semantic considerations (via an exclusionary expressivism), however well semantic nominalization nominalism *fits with* ontological nominalism, the semantic thesis is not in the right shape to provide an *argument* for the ontological one—as Sellars in effect claims that it is in the passage from *NS* I quoted above. Even if the semantic claim that transcategorial nominalizations are not genuine (referring) singular terms is accepted, that in no way entails that only what *can* be so referred to exists in the real world. Such an ontological stipulation is at most *compatible* with the semantic commitment. So I do not think that there is an argument from Sellars’s metalinguistic pragmatic and semantic nominalization nominalism to his ontological nominalism.

Nor can I see that the scientific realism epitomized in Sellars's *scientia mensura* passage—"In the dimension of describing and explaining the world, science is the measure of all things, of what is that it is, and of what is not that it is not,"—yields an argument for reistic ontological nominalism.<sup>186</sup> That is, it would not help to restrict what exists in "the world in a narrow sense" to what can be *described*. The descriptive language of science is just as much up for alternative interpretations, nominalistic and otherwise, as ordinary language. If all that is right, then we should see Sellars's commitment to a reistic ontological nominalism of the sort epitomized by Kotarbiński (before his pan-somatist turn) as rock-bottom, not derived from or supported by other commitments. His metalinguistic expressivism about transcategorial nominalizations should be understood as aimed at showing that one *need* not countenance universals and propositions to understand the use of the expressions that ostensibly refer to them.

I conclude that Sellars has introduced and deployed the metalinguistic machinery of dot-quotes, distributive singular terms, and Jumblese to offer a sophisticated account of a distinctive metalinguistic role that transcategorial nominalizations and their associated common nouns play. That account, though, operates primarily at the level of *pragmatics*: as part of a theory of the *use* of these expressions. He has *not* thereby put himself in a position to be entitled to draw nominalistic *semantic* or *ontological* conclusions from the identification of that distinctive expressive role. In the absence of a fuller analysis of this case, we should no more draw that conclusion from Sellars's expressivist analysis of the use of property-terms than we should from his expressivist account of the use of alethic modal vocabulary.

END

---

<sup>186</sup> *EPM* §41.

