Negotiating Constitutivity: A Pragmatist Account of Interpretive Coordination

DISSEYRATI6N

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in
the Graduate School of The Ohio State University

By

Michael David Miller, B.A.

Graduate Program in Philosophy

The Ohio State University

2014

Dissertation Committee:

Robert Kraut, Advisor

Kevin Scharp

Stewart Shapiro
Abstract

I begin by tracing the historical origins of the analytic-synthetic distinction and isolating three distinct philosophical roles—metaphysical, epistemological, and interpretive—that analytic sentences have traditionally been expected to play. In its metaphysical role, analyticity is supposed to explain the nature and source of necessary truth; in its epistemological role, analyticity is supposed to provide a ground for a priori knowledge; and in its interpretive role, analyticity is supposed to explain competent speakers’ understanding of linguistic expressions. I review the most important criticisms of analyticity’s alleged ability do this explanatory work, and argue that the three most well worked-out recent attempts to revive a notion of analyticity fail to yield a notion that is able to play even one of analyticity’s three philosophical roles.

Although I take the metaphysical and epistemological roles of analyticity to have been decisively undermined by Quine and others, I argue that allegedly analytic sentences very often do play something like the interpretive role. Such sentences are treated as pro tanto interpretive guidelines by which assessments of linguistic understanding and judgments about sameness of meaning are made. However, I reject the assumption that such sentences are analytic truths which must be believed in order to understand the expressions they contain, and urge instead that we characterize such sentences in terms of the non-factive notion of constitutivity.
I develop a pragmatic account of interpretive coordination according to which the standards of usage governing shared linguistic expressions are rationally negotiated among the interlocutors in a discourse interaction. The interpretive standards thus established provide the interlocutors with a way to distinguish between changes of meaning and changes of belief, and thus between verbal and substantive disputes. The process by which these interpretive standards are rationally negotiated is explained in terms of behavioral dispositions I call constitutivity commitments.

I argue that our ability to align our constitutivity commitments is what makes communication (i.e., the sharing of information) possible, and helps explain the emergence of the communal usage regularities characteristic of human linguistic practice. Very roughly, constitutivity commitments are practical interpretive dispositions to react to the denials of certain sentences in characteristic ways. Typically, this reaction involves deployment of the sort of intensional vocabulary W. V. Quine famously raises skepticism about in “Two Dogmas of Empiricism.” As such, I argue that the pragmatic function of both meaning-talk and essence-talk is to express our constitutivity commitments, but I avoid Frege-Geach worries by rejecting semantic non-factualism in favor of a thoroughgoing deflationism about semantic representation.
Dedicated to my parents, Aidan and Myra Miller.
Acknowledgments

Thanks first and foremost to my advisor, Robert Kraut. When we weren’t chatting about our cats, vintage tube amps, or how to axiomatize the work of Grand Funk Railroad, Robert was showing me how to think clearly and carefully about the deepest and most difficult issues in philosophy. I will be forever grateful for his invaluable wisdom, guidance, encouragement, and friendship. I want to thank Kevin Scharp, whose philosophical influence on me has been profound, and whose generous feedback at every stage of the project has improved this dissertation beyond measure. Thanks also to Stewart Shapiro for his many insightful suggestions, challenging feedback, and ongoing support. I am indebted to Craige Roberts for helpful discussion and valuable advice on how to integrate constitutivity commitments into formal pragmatics. Over the years, I have had the honor of learning from many exceptional teachers, but I am especially grateful to William Taschek, Lisa Shabel, Neil Tennant, Alan Silverman, and Abe Roth. I am also grateful to the dissertation seminar participants who read and commented on my work, especially Wesley Cray, Dai Heide, Owen King, and Alison Kerr. Finally, I want to thank my perpetually supportive family: my parents Aidan and Myra, my brothers Daniel and Adam, and my loving wife Courtney.
Vita

July 12, 1983............................................................... Born, Dearborn Michigan

2005........................................................................... B.A., Philosophy
          University of Michigan – Flint

2005-2013............................................................... Graduate Teaching Associate,
          The Ohio State University

Fields of Study

Major Field: Philosophy
# Table of Contents

Abstract ............................................................................................................................................. ii

Dedication ........................................................................................................................................ iv

Acknowledgments ........................................................................................................................... v

Vita..................................................................................................................................................... vi

Chapters:

1. A Brief History of Analyticity ........................................................................................................ 1
   1.1 Introduction .............................................................................................................................. 1
   1.2 The Emergence of Analyticity: Locke, Hume, and Kant......................................................... 3
   1.3 Clarifying Kant’s Distinction: Frege and Carnap................................................................. 9
   1.4 Variations on a Theme: Ayer and Waismann........................................................................ 15
   1.5 The Three Philosophical Roles of Analyticity..................................................................... 21
   1.6 Is Analyticity a Legitimate Theoretical Notion? ................................................................. 24
   1.7 Can Conventional Stipulation Explain Necessary Truth?................................................... 26
   1.8 Can Implicit Definition Ground *A Priori* Knowledge?.................................................... 28
   1.9 Can Analyticity Explain Linguistic Understanding?............................................................. 32
   1.10 Concluding Remarks.............................................................................................................. 36

2. Three Recent Accounts of Analyticity: A Critical Survey........................................................ 38
   2.1 Introduction .............................................................................................................................. 38
   2.2 Boghossian on “Epistemic Analyticity”............................................................................... 38
   2.3 Russell on “Truth in Virtue of Meaning”.............................................................................. 46
   2.4 Juhl and Loomis on “Analyticity**”.................................................................................... 53
Chapter 1

A Brief History of Analyticity

1.1 Introduction

The analytic-synthetic distinction has an impressive philosophical pedigree. It was first explicitly articulated by Kant, who had been influenced by Hume’s distinction between Relations of Ideas and Matters of Fact. Hume and Kant were both concerned with the nature of necessity and a priority, and these distinctions were put to work in the service of fundamental metaphysical and epistemological concerns. Similar concerns motivated Frege and Carnap to take up and refine Kant’s distinction, demonstrating the crucial role it played in the development of the analytic tradition. It was not until the middle of the last century, due most significantly to the work of Quine, that skepticism began to take root about the philosophical viability of the venerable analytic-synthetic distinction. In the decades that followed the publication of Quine’s classic paper “Two Dogmas of Empiricism,” many became convinced that the philosophical legitimacy of appeals to the analytic-synthetic distinction had been decisively undermined. In recent years, however, there has been renewed interest in the notion of analyticity, including a number of attempts to re-engineer it or introduce some closely related replacement notion.

An analytic statement is supposed to be true by virtue of meaning alone, while the truth of a synthetic statement is supposed to depend both on meaning and on the way the
world is. Because analytic sentences are true by virtue of meaning alone, anyone who understands them can come to have *a priori* knowledge of their truth just by reflecting on the meanings of their component terms. Since analytic sentences define or constitute the meaning of the expressions they contain, competent language-users are entitled to hold them true no matter what the world turns out to be like. By contrast, the truth of synthetic statements can only be established on the basis of empirical evidence. Once a person has learned a language, she appears to be able to acquire analytic knowledge simply by reflecting on the meanings of her words; and so long as those meanings are held fixed, it seems that these analytic statements cannot possibly be false—they hold true of necessity.

Two of the most important questions that animate the analytic tradition are these: (1) *how is a priori knowledge possible?* and (2) *what is the nature and source of necessary truth?* For those wishing to address these questions in a naturalistic spirit, an explanatory strategy that appeals to linguistic meaning can seem to be an attractive alternative to the postulation of a mysterious faculty of rational intuition by which the mind somehow secures *a priori* access to objective modal features of the world. If analytic sentences are true by virtue of meaning alone, perhaps the notion of analyticity is the key to explaining both *a priori* and necessity. The success of such a strategy, however, depends on whether the notion of meaning can be rendered naturalistically acceptable. Moreover, even if the naturalist is able to domesticate meaning, there is still the further question of whether semantic notions are really fit to bear the explanatory burden assigned to them.

In this chapter, I do three things. First, I outline the historical origins of the analytic-synthetic distinction. Second, I discuss the three essential philosophical roles—metaphysical, epistemological, and interpretive—that analytic statements have historically been expected to
play. Finally, I discuss the most important reasons for skepticism about analyticity and then briefly foreshadow my own positive proposal.

1.2 The Emergence of Analyticity: Locke, Hume, and Kant

John Locke appears to endorse a version of what would later be labeled the analytic-synthetic distinction, though he does not call it by that name. Locke calls identity claims (e.g., ‘A soul is a soul’; ‘A vortex is a vortex’) and claims in which “a part of any complex idea is predicated of the name of the whole” (e.g., ‘Lead is a metal’; ‘All gold is fusible’; ‘A palfrey is an ambling horse’) _trifling propositions_. He characterizes such propositions as certainly (though trivially) true, involved in definitions, not apt to extend substantive knowledge, and points out their use in teaching the significance of the words they contain. The examples of trifling propositions given by Locke will later be regarded as paradigm cases of analyticity by Kant and others.

According to Locke, trifling propositions form a special class and are to be contrasted with the substantive, non-trivial claims that may genuinely extend our knowledge of non-linguistic reality. Locke notes that identity claims “wherein the same term is affirmed of itself” are not particularly useful for linguistic instruction, and makes clear that it is the other sort of trifling proposition (“in which a part of any complex idea is predicated of the whole”) which may usefully serve this purpose. This class of trifling propositions specify “a part of the definition of the word defined” (§4), and may thus be informative to a novice who is unaware of the complete definition. But to anyone who already understands the terms in these propositions, asserting them is unnecessary and redundant, which is why

---

1 See Locke (1690), chapter VIII.
Locke says that “he trifles with words who makes such a proposition” (§7). However, in a situation in which someone “declares himself not to understand” a certain word, an otherwise “trifling” proposition may be used to teach him “the signification of that word, and the use of that sign” (§7).

Locke admits, however, that there may be necessary consequences of certain propositions that are not strictly “contained in” a particular complex idea. In this, he anticipates Kant. Locke writes:

We can know then the truth of two sorts of propositions with perfect certainty. The one is, of those trifling propositions which have a certainty in them, but it is only a verbal certainty, but not instructive. And, secondly, we can know the truth, and so may be certain in propositions, which affirm something of another, which is a necessary consequence of its precise complex idea, but not contained in it: as that the external angle of all triangles is bigger than either of the opposite internal angles. Which relation of the outward angle to either of the opposite internal angles, making no part of the complex idea signified by the name triangle, this is a real truth, and conveys with it real knowledge (§8).

Locke thus holds that our complex concepts have a core “nominal essence” which contains other concepts as parts, and that possessing such a concept consists in (verbal) knowledge of those core components. There may be additional (non-verbal) knowledge to be gained about certain things that we have complex concepts of (e.g., triangles), but Locke seems to limit this to mathematical or geometric concepts, rejecting as empty claims such as ‘All men have a notion of God’ and ‘All men are cast into sleep by opium’ on the grounds that having a notion of God or being cast into sleep by opium are not among the essential components of the concept man (§6, §9).

David Hume builds on the many of the themes explored by Locke when he proposes his famous distinction between Matters of Fact and Relations of Ideas. Hume is concerned to develop a science of human nature which is thoroughly empirical and free

---

2 See Hume (1748), §IV, pt. I.
from dubious metaphysical speculation not grounded in experience. Hume uses the term ‘perception’ to denote any mental item of which we are consciously aware, and he divides these items into two sub-classes:

Here therefore we may divide all the perceptions of the mind into two classes or species, which are distinguished by their different degrees of force and vivacity. The less forcible and lively are commonly denominated Thoughts or Ideas. The other species want a name in our language... Let us... call them Impressions, employing that word in a sense somewhat different from the usual. By the term impression, then, I mean all our more lively perceptions, when we hear, or see, or love, or hate, or desire, or will (§II).

Impressions are either directly produced by sensory stimulation (outward sentiment) or by introspective awareness of one’s inner mental states (inward sentiment). Thoughts or Ideas are mental items formed by copying impressions, and lose some “liveliness” or “vivacity” in the copying process.

According to Hume, ideas can be simple (i.e., not built up out of other ideas) or complex (i.e., composed of other component ideas). Complex ideas are constructed by augmenting and recombining various simple ideas originally derived from sensory input. Hume writes:

When we think of a golden mountain, we only join two consistent ideas, gold and mountain, with which we were formerly acquainted. A virtuous horse we can conceive; because, from our own feeling, we can conceive virtue; and this we may unite to the figure of a horse, which is an animal familiar to us. In short, all the materials of thinking are derived either from our outward or inward sentiment: the mixture and composition of these belongs alone to the mind and will. Or, to express myself in philosophical language, all our ideas or more feeble perceptions are copies of our impressions or more lively ones. (§II)

The final sentence of this passage is a statement of Hume’s Copy Principle, which he uses to distinguish genuinely intelligible discourse from empty metaphysical claims. Proper application of the Copy Principle is claimed to “render every dispute equally intelligible, and banish all that jargon, which has so long taken possession of metaphysical reasonings, and drawn disgrace upon them” (§II).
After defining these basic concepts and principles, Hume goes on to draw a

distinction that sets the groundwork for what will become, in Kant’s hands, the analytic-
synthetic distinction. He writes:

All of the objects of human reason or enquiry may naturally be divided into two kinds, to
wit, Relations of Ideas and Matters of Fact. Of the first kind are the sciences of Geometry,
Algebra, and Arithmetic; and in short, every affirmation which is either intuitively or
demonstratively certain. That the square of the hypotenuse is equal to the square of the two sides, is a
proposition which expresses a relation between these figures. That three times five is equal to the
half of thirty, expresses a relation between these numbers. Propositions of this kind are
discoverable by the mere operations of thought, without dependence on what is anywhere
existent in the universe. … Matters of fact, which are the second objects of human reason,
are not ascertained in the same manner; nor is our evidence of their truth, however great, of
a like nature with the forgoing. The contrary of every matter of fact is still possible; because
it can never imply a contradiction. That the sun will not rise to-morrow is no less intelligible a
proposition, and implies no more contradiction than the affirmation, that it will rise. We
should in vain, therefore, attempt to demonstrate its falsehood. Were it demonstrably false, it
would imply a contradiction, and could never be distinctly conceived by the mind.” (§IV)

Here, Hume cites mathematical statements as paradigm examples of Relations of Ideas,
noting several of their definitive attributes: that they are intuitively or demonstrably certain,
knowable by ratiocination alone without need of sensory input, that they do not depend on
the existence of any external things, and that their negations are contradictory and thus
inconceivable. Any truth that lacks the above characteristics counts as a Matter of Fact.

Immanuel Kant was deeply impressed by Hume’s work, famously remarking that
Hume’s critique of speculative metaphysics awoke him from his “dogmatic slumber.”3 Prior
to having read Hume, Kant was a committed Rationalist who believed that substantive
metaphysical knowledge about the nature of the world and its objects could be gotten via
pure a priori contemplation. His mature views, tempered by Hume’s empiricist influence,
were much more subtle and innovative.

3 See Kant (1783), introduction.
Kant was the first to explicitly articulate the analytic-synthetic distinction, and it is from him that we get the terms ‘analytic’ and ‘synthetic’. Kant’s account of the analytic-synthetic distinction is clearly informed by Hume’s account of the distinction between Matters of Fact and Relations of Ideas. However, Kant’s distinction is more fine-grained, leaving room for a variety of a priori knowledge which Hume’s distinction cannot countenance. Kant’s first discussion of the analytic-synthetic distinction appears in the introduction to the first edition of the *Critique of Pure Reason*:

In all judgments in which the relation of a subject to the predicate is thought (if I only consider affirmative judgments, since the application to negative ones is easy) this relation is possible in two different ways. Either the predicate \( B \) belongs to the subject \( A \) as something that is (covertly) contained in this concept \( A \); or \( B \) lies entirely outside the concept \( A \), though to be sure it stands in connection with it. In the first case, I call the judgment analytic, in the second synthetic. (A:6-7)

Later, in his *Prolegomena to Any Future Metaphysics*, Kant gives the following further gloss: “Analytic judgments say nothing in the predicate except what was actually thought already in the concept of the subject, though not so clearly nor with the same consciousness.” Kant’s basic idea, then, is that analyticity is a matter of the “covert containment” of one concept within another: A judgment is analytic if the concept expressed by the predicate term is wholly contained within the concept expressed by the subject term—otherwise, the judgment is synthetic.

Since they describe containment relations between fully determinate concepts, the denial of an analytic claim will always be a manifest contradiction. Thus, on Kant’s account, analytic judgments are necessarily true. Kant is also quite explicit that analytic judgments are knowable a priori. In the second edition of the *Critique*, Kant writes:

---

4 Kant (1781/1787).
5 Kant (1783), §2.
That a body is extended is a proposition that is established \textit{a priori}, and is not a judgment of experience. For before I go to experience, I already have all the conditions for my judgment in the concept, from which I merely draw out the predicate in accordance with the principle of contradiction, and can thereby at the same time become conscious of the necessity of the judgment, which experience could never teach me. (B12)

And then in the \textit{Prolegomena}:

All analytic judgments rest entirely on the principle of contradiction and are by their nature \textit{a priori} cognitions, whether the concepts that serve for their material be empirical or not. For since the predicate of an affirmative analytic judgment is already thought beforehand in the concept of the subject, it cannot be denied of that subject without contradiction; exactly so is its opposite necessarily denied of the subject in an analytic, but negative, judgment, and indeed also according to the principle of contradiction. (§2)

Because “conceptual containment” may be “covert” in various cases, Kant allows that analytic judgments need not be obvious or unsurprising. But in principle, analytic judgments can be known, and their necessity recognized, just by reflecting carefully on the nature of the concepts which feature in the judgment. A consequence of this is that anyone who (fully) possesses the subject-concept of an analytic judgment must thereby also possess the predicate-concept, and only someone who did not (fully) possess the subject-concept could sincerely deny the analytic judgment in question.

Kant’s way of drawing the analytic-synthetic distinction deliberately leaves room for the possibility of \textit{synthetic, yet a priori} judgments—i.e., substantive claims about the external world which, though not analytic, are still knowable \textit{a priori} because of the way our human minds must be structured in order to have any experience at all. Where Hume assimilates the truths of arithmetic and geometry into Relations of Ideas, Kant takes arithmetic propositions (e.g., $7 + 5 = 12$) and geometrical propositions (e.g., the axioms and theorems of Euclidean geometry) to be paradigm cases of the synthetic \textit{a priori}.

This move, however, ends up making serious trouble for Kant’s overall philosophical system. For in the wake of Einstein’s general theory of relativity, which understands
gravitation in terms of spacetime warpage produced by massive bodies, physicists came to realize that the geometry of actual space is non-Euclidean. In light of this empirical discovery, we must admit that, pace Kant, the propositions of Euclidean theory fail to get the world right, and a fortiori, cannot be candidates for a priori knowledge. This result and others like it drove many naturalistically-minded philosophers to become deeply skeptical of Kantian appeals to so-called “synthetic a priori” knowledge. Moreover, even if there is such a thing as the synthetic a priori, Kant gives us no reason to think that human beings can reliably distinguish between genuinely synthetic a priori propositions and those that merely appear to be so. Unfortunately for Kant, being entitled to firmly believe Euclidean theory, given everything else you take yourself to know, turns out to be phenomenologically indistinguishable from having genuine synthetic a priori knowledge that Euclidean theory is true.

1.3 Clarifying Kant's Distinction: Frege and Carnap

Gottlob Frege’s principle philosophical goal throughout his writings is to provide a foundation for arithmetic, which he hopes to do by showing that the all the truths of arithmetic can in principle be derived from the laws of logic. It is in the service of this goal that Frege develops his version of the analytic-synthetic distinction. He takes issue with various aspects of Kant’s account of the distinction, and seeks to offer an improved and clarified version of it. In Die Grundlagen der Arithmetik, Frege writes:

Kant obviously underestimated the value of analytic judgments—no doubt as a result of defining the concept too narrowly, although the broader concept used here does appear to have been in his mind. On the basis of his definition, the division into analytic and synthetic judgments is not exhaustive. He is thinking of the case of universal affirmative judgment. Here one can speak of a subject-concept and ask—according to the definition—whether the predicate concept is contained in it. But what if the subject is an individual object? What if
the question concerns an existential judgment? Here there can be no talk at all of a subject-concept in Kant’s sense. Kant seems to think of a concept as defined by a conjunction of marks; but this is one of the least fruitful ways of forming concepts.\(^6\)

Armed with his new system of quantificational logic, which could handle much more than just affirmative categorical subject-predicate judgments, Frege attempts to explain our \textit{a priori} knowledge of arithmetic by showing that the axioms of arithmetic are analytic. Like Kant, Frege accepts that if a judgment is knowable \textit{a priori}, as analytic judgments are, then it cannot fail to be true. For Frege, “An \textit{a priori} error is thus just as much an absurdity as, say, a blue concept.”\(^7\)

Frege objects to Kant’s murky “covert containment” idea, which he feared might be unacceptably psychologistic.\(^8\) Surely the question of whether a certain proposition is analytic is an objective matter, not a question about anyone’s psychology. Frege thus defines a proposition as analytic just in case it is \textit{provable using only general logical laws and definitions}. In support of this proposal, he writes:

\begin{quote}
In this way the question [of whether a proposition is analytic] is removed from the domain of psychology and assigned to that of mathematics, if it concerns a mathematical truth. It now depends on finding a proof and following it back to the primitive truths. If, on the way, only general logical laws and definitions are encountered, then the truth is analytic, assuming that propositions on which the admissibility of any definition rests are also taken into account. If it is not possible to provide a proof, however, without using truths that are not of a general logical nature, but belong instead to the domain of a particular science, then the proposition is synthetic.\(^9\)
\end{quote}

This new formulation of the analytic-synthetic distinction allowed Frege to argue that, \textit{pace} Kant, the truths of arithmetic are all analytic because they are reducible to logical truths.

\(^6\) Frege (1884), §88.
\(^7\) Ibid., §3.
\(^8\) See Katz (1966), (1997) for an account of analyticity in terms of “containment relations” between the senses of various expressions. Katz takes senses to be “in the head” and attempts to avoid externalist arguments against narrow content by rejecting the crucial Fregean thesis that sense determines reference. See Cohen (2000) and Linsky (1970) for powerful criticisms of Katz’s project.
\(^9\) Frege (1884), §3.
Frege thus had what he considered to be a powerful explanation of the necessity and *a priori* of arithmetic. Since arithmetical truths are all analytic—that is, they may be proved using only axiomatic definitions and logical laws—arithmetical reduces to logic. And, on the assumption that logic is necessary and *a priori*, so is arithmetic. Of course, one important worry about this move is that even if the proposed reduction works (Russell’s paradox notwithstanding), Frege has not yet explained the *a priori* and necessity of logic itself, and thus the story remains importantly incomplete. Frege, for his part, did not think the further explanatory demand could be met, and was content to simply treat logical laws as primitively valid. Many of Frege’s philosophical descendants, however, would not be satisfied with such primitivism, and would strive instead for a more complete account of the *a priori* which included knowledge of logic.

The next big development in the history of analyticity comes when the logical positivists—and in particular, Rudolf Carnap—get a hold of it. Carnap is interested in developing *explications* (i.e., improved, precisified replacements) for what he takes to be untidy, inexact philosophical concepts. He takes the notion of analyticity to be an excellent candidate for explication, and hopes to engineer a precise version of the analytic-synthetic distinction that can be used to do what Frege’s version could not—namely, account for the *a priori* and necessity of logic.\(^\text{10}\) Following Wittgenstein’s suggestion in the *Tractatus Logico-Philosophicus*, Carnap conceives of the truths of logic as tautologies which reveal the inferential relations between various propositions but do not say anything about the

---

\(^{10}\) See Wild and Coblitz (1948) for an early (i.e., pre-“Two Dogmas”) critical discussion of both the positivist version of the analytic-synthetic distinction and Kant’s original version. See Beck (1949) for a Kantian reply.
empirical world. He thinks that mathematics lacks empirical content in the same way, and he aims to provide an account of our knowledge of both.

Carnap’s philosophy went through two different phases: an earlier “syntactic” period and a later “semantic” period. The shift between these periods occurs when Carnap becomes aware of Alfred Tarski’s work on truth, which convinces him that semantic notions can be made sufficiently precise for rigorous theoretical purposes. The focus here will be on Carnap’s semantic version of analyticity, but first it is worth briefly considering a fundamental problem with Carnap’s earlier non-semantic version.

During his syntactic period (i.e., in *Logical Syntax of Language*), Carnap attempts to explicate analyticity and synonymy as purely formal or syntactic notions. The problem with Carnap’s early approach is that to treat analyticity and synonymy as purely formal properties is to sever the connection between sense and reference, and thus sever the connection between a sentence’s being *analytic* and its being *true*. For if ‘Sentence S is analytic’ attributes a purely formal non-semantic property to S, then ‘S is true’ does not immediately follow from ‘S is analytic’. However, because analyticity is an essentially factive notion, any viable account of analyticity must accommodate the fact that ‘Sentence S is analytic’ immediately entails ‘S is true’. Indeed, analytic sentences can’t be knowable *a priori* unless they are true. The upshot is that analyticity cannot be understood as a purely formal, non-semantic notion. Thus, arguably, Carnap’s earlier non-semantic notion is not really a species of analyticity at all.

---

11 See Wittgenstein (1922), §6.1-6.11.  
13 See Carnap (1934/1937).  
In his later semantic period, (i.e., in *Meaning and Necessity*), Carnap explicates analyticity in terms of the notion of *L-truth*, such that a sentence is analytic (i.e., L-true) just in case “the semantical rules of the system suffice for establishing its truth.”\(^{15}\) Carnap goes on to explicate the synthetic as follows:

A sentence is called *L-determinate* if it is either L-true or L-false; otherwise it is called L-indeterminate or *factual*. The latter concept is an explicatum for what Kant called synthetic judgments. A sentence is called *F-true* if it is true but not L-true; F-truth is an explicatum for what is known as factual or synthetic or contingent truth.\(^{16}\)

To further sharpen the notion, Carnap explains that analytic sentences are those that “hold in all state-descriptions”—where the notion of a *state-description* is Carnap’s intended explication of Leibniz’ *possible worlds* or Wittgenstein’s *possible states of affairs.*\(^{17}\) A state-description (for a language \(L\)) is an exhaustive assignment of truth-values to all the literals (i.e., atomic sentences or their negations) in \(L\). By the “semantical rules” (for a language \(L\)), Carnap means the enumerated postulates that stipulate how to assign truth-values to the non-literals in a given state-description of \(L\). Thus, for Carnap, given a language \(L\), the term ‘analytic’ applies to (a) the semantical rules of \(L\); and (b) any logical consequences of the semantical rules of \(L\).

With these mechanisms in place, Carnap attempts to use the notion of analyticity to accomplish a much more ambitious set of philosophical goals than Frege had attempted. As noted above, even if Frege’s reduction of arithmetic to logic had worked, we still would have lacked an adequate account of the *a priori* and necessary status of logic itself. Carnap sought to remedy this by conjoining his views on analyticity with a doctrine of conventionalism about the truths of logic. According to Carnap, which language we adopt and use is purely a

\(^{16}\) Ibid.
\(^{17}\) Ibid., p. 9.
matter of convention. The choice of which language to use is to be decided solely on pragmatic grounds: if adopting the language in question would be useful to us, no further justification is needed for doing so. Moreover, by conventionally adopting a language, we thereby conventionally adopt the semantical rules which are stipulated to govern the use of that language. And since L-truth, or analyticity, is determined by the semantical rules stipulated to govern the language we are using, which sentences turn out to be L-true is purely a matter of convention. We are free to construct languages—or, in Carnap’s terminology, linguistic frameworks—in any way we like, so long as we explicitly state the semantical rules that govern the proper application of the expressions in our proposed framework. To adopt a linguistic framework just is to adopt a certain set of semantical rules which specify the proper use of the framework’s vocabulary items.

For Carnap, once we have adopted a certain framework, there can be no further question as to which semantical rules are the “correct” ones. From within the framework, the semantical rules are trivially correct, since they specify the very standards by which assessments of correctness are to be made. From outside the framework, the question of correctness simply makes no sense. For to exit the framework is to either (a) adopt some other framework with standards of its own, or (b) attempt to stand outside of all frameworks and operate without any standard of correctness. But in the absence of any such standard, questions about the correctness of a framework’s rules are unanswerable pseudo-questions.

Carnap takes the above picture to constitute an adequate explanation of both the necessity and a priority of logic. Being analytic, logical truths are rule-specifying, framework-constituting principles that do not represent worldly states of affairs and thus can never be

\footnote{See Carnap (1950).}
empirically disconfirmed. We are justified in holding them true not on empirical grounds, but because they are among the semantical rules which we have stipulated shall govern the proper use of our language.

1.4 Variations on a Theme: Ayer and Waismann

In *Language, Truth & Logic*, A.J. Ayer proclaims that philosophy is “a special branch of knowledge” whose function is “essentially analytic.”19 To engage in philosophical analysis, so conceived, is to offer up a definition of a word “by showing how the sentences in which it significantly occurs can be translated into equivalent sentences, which contain neither the *definiendum* itself, nor any of its synonyms.”20 The statement of such a translation expresses a complex analytic proposition. Ayer also endorses a “criterion of verifiability” according to which the empirical content of each well-formed indicative sentence is to be identified with the set of actual or possible observations that would provide evidence for or against it.21 Any meaningful sentence that lacks verification conditions is a tautology, and any non-analytic sentence that lacks verification conditions is a meaningless nonsense (paradigm examples include metaphysical and theological statements).

On Ayer’s view, the special knowledge gained by philosophical analysis is distinctively non-empirical and thus cannot conflict with our scientific knowledge:

In other words, the propositions of philosophy are not factual, but linguistic in character—that is, they do not describe the behavior of physical, or even mental, objects; they express definitions, or the formal consequences of definitions. Accordingly, we may say that philosophy is a department of logic. For we shall see that the characteristic mark of a purely logical enquiry is that it is concerned with the formal consequences of our definitions and not with questions of empirical fact.22

---

19 Ayer (1936), pp. 51-52.
20 Ibid., p. 60.
21 Ibid., p. 35.
22 Ibid., p. 57.
In this, he echoes Carnap, who conceives of typical philosophical claims (about, e.g., the nature of time, space, or number) as “quasi-syntactical” or “pseudo-object” sentences which are systematically, but misleadingly, formulated in the “material mode” of speech.\(^{23}\) Despite their object-level formulation, such claims are, as Alberto Coffa puts it, “not about the world but about the structure of an object language in which, in turn, we talk about the world.”\(^{24}\)

Likewise, according to Ayer, analytic propositions “do not make any assertion about the empirical world” but instead “simply record our determination to use words in a certain fashion.”\(^{25}\) Presumably, then, Ayer and Carnap are committed to the idea that the languages we employ can be engineered so that they have a determinate inferential structure, and that when we make analytic claims we are describing such a structure.\(^{26}\)

Of all the thinkers associated with the Logical Positivist movement, Friedrich Waismann is perhaps the most widely and unfairly neglected. This is unfortunate; for despite being largely ignored, Waismann’s work on meaning and analyticity is extremely subtle and insightful. The bulk of this work appears in his 1945 paper “Verifiability” and in the six-part series “Analytic-Synthetic” published between 1949 and 1953.\(^{27}\)

Before we explore Waismann’s views on analyticity, let’s consider first what he has to say about meaning and verification. According to Waismann, when we state the verification conditions of a sentence, what we are doing is endorsing the legitimacy of a certain inference

---


\(^{25}\) Ayer (1936), p. 84.

\(^{26}\) This amounts to what Coffa (1991) calls a “second-level semantic factualism” according to which “there is a fact of the matter concerning the difference between the stage at which we produce the semantic machinery involved in communication and the stage at which we are finally communicating—or if you will, the analytic-synthetic distinction” (p. 322).

\(^{27}\) Also relevant is Waismann (1946), which contains sketches of some of the ideas and themes explored at greater length in the “Analytic-Synthetic” series.
pattern from a particular perspective. This is not a factual description of how the statement is used by some community or other, but rather a prescriptive claim regarding how, by the speaker’s lights, it ought to be used. He notes that in ordinary contexts we typically take these inference patterns for granted and don’t bother making them explicit. Thus, according to Waismann:

The question of verification arises only when we come across a new sort of combination of words. If, for instance, someone were to tell us that he owned a dog that was able to think, we should at first not quite understand what he was talking about and would ask him some further questions. Suppose he described to us in detail the dog’s behaviour in certain circumstances, then we should say ‘Ah, now I understand you, that’s what you call thinking’. There is no need to inquire into the verification of such sentences as ‘The dog barks’, ‘He runs’, ‘He is playful’, and so on, as the words are then used as we may say in their normal way. But when we say ‘The dog thinks’ we create a new context, we step outside the boundaries of common speech, and then the question arises as to what is meant by such a word series. In such cases explaining the verification is explaining the meaning, and changing the verification is changing the meaning.

The idea is that when we ask questions about what a certain sentence means, what we are after is a sense of how it links up inferentially with other sentences, and in particular, what would count as evidence for it.

But although Waismann agrees that the meaning is connected to verification, he does not think that sentences are content-equivalent to their verification conditions because the latter need not entail the former. He thus rejects reductive phenomenalism: “[A] material object statement, though it is connected with sense-datum statements, is not just an abbreviation for them, rather it has a logical status of its own, and is not equivalent to any

---

28 Waismann (1945), p. 117.
29 Ibid., p. 118.
30 Later on, however, Waismann seems more hostile to the notion of meaning. In particular, see Waismann (1951a) where he says that “there is no sharp line which separates those uses which, as one would say, are characteristic of the concept, from those that are not,” and concludes that “to speak of the ‘meaning’ of a word and to ask whether it has, or has not changed in meaning, is to operate with too blurred an expression” (pp. 50-53).
truth-function of the latter ones.” Waismann goes on to give a novel diagnosis of the failure reductionist programs in general: such projects are doomed to fail because of what he calls the open texture of most of our concepts.

An expression that exhibits open texture is “not delimited in all possible directions” in the sense that “we can never exclude altogether the possibility of some unforeseen situation arising in which we shall have to modify the definition.” According to Waismann,

We introduce a concept and limit it in some directions; for instance, we define gold in contrast to some other metals such as alloys. This suffices for our present needs, and we do not probe any farther. We tend to overlook the fact that there are always other directions in which the concept has not been defined.

In this way, the definitions of open terms are essentially incomplete. By contrast, a complete definition would specify “an exhaustive list of all the circumstances in which the term is to be used” which conclusively “anticipates and settles once and for all every possible question of usage.” Complete definitions are only possible for the terms of logic and mathematics, not for empirical terms. Because they can be completely defined in this way, mathematical and logical concepts exhibit closed texture.

For Waismann, the point of giving a definition is not to describe the predominant usage of a word but to take a stand on how that word is to be used. He emphasizes the role of definitions in linguistic instruction and the fact that speakers often appeal to definitions when called upon to defend the legitimacy of their usage. Moreover, on his view, a principle’s status as a definition is not static but dynamic. Nothing has definitional status

32 Ibid., p. 119. See also Waismann (1946), pp. 224-229.
33 Waismann (1945), p. 120.
34 Ibid.
36 Ibid., p. 123. See also Waismann (1951b), p. 118.
essentially or immutably because what counts as a definition is always subject to change as we acquire new information.

In light of this dynamic conception of definitional status, Waismann holds that what is most important about our practice of giving definitions has to do with the inferential transitions we are thereby able to explicitly endorse. For example, suppose I take the sentence ‘A planet is a heavenly body revolving around the sun’ to count as a definition of the term ‘planet’. This allows me to replace the word ‘planet’ with the phrase ‘a heavenly body revolving around the sun’ thus making it possible to transform the original statement into a logical truth. According to Waismann, such a definition can “be construed as a substitution license which gives permission to interchange these two locutions.” In this way, definitions are similar to claims of sentence-synonymy, the chief difference being that the former permit term (or phrase) substitution where the latter permit the substitution of whole sentences.

On the basis of these considerations, Waismann arrives at the following provisional characterization of analyticity: “A statement is analytic if it can, by means of mere definitions, logical and, further, idiomatic (linguistic) operators, be turned into a truth of logic.” He notes that this transformation process is not purely logical, since it depends on definitions and “idiomatic” operators (i.e., statements of sentence-synonymy). The acceptability of definitional and idiomatic substitutions is not decided by the logical axioms but depends instead on how the relevant words or sentences are (properly) used in the speaker’s language. Given the open texture of natural language expressions, however, the status of particular definitional and idiomatic substitutions is always subject to change as we

38 Ibid., p. 36.
confront previously unforeseen circumstances and are forced to make decisions about how to apply our terms going forward. This leads Waismann to the following conclusion:

[I]t is significant that we do not only “find out” that a given statement is analytic; we more often precisify the use of language, chart the logical force of an expression, by declaring such-and-such a statement to be analytic. If analytic was as fixed and settled a term as, say, ‘tautology’ is, this would be hard to understand: can I e.g., by decree appoint a given statement to the rank of tautology? It is precisely because, in the case of ‘analytic’, the boundary is left open somewhat that, in a special instance, we may, or may not, recognize a statement as analytic.\footnote{Waismann (1950), p. 25.}

So although there is no room for dispute about which statements are to count as tautologies, there is room for dispute about what counts as analytic on Waismann’s view. The reason for this is that a statement’s analytic status depends crucially on which definitional and idiomatic substitutions we are prepared to endorse, and since the definitions of open-textured concepts are fundamentally incomplete and perpetually unsettled, what counts as an analytic statement is not fixed either. As a result, analytic statements are “on the borderline between necessary and contingent, the \textit{a priori} and the empirical.”\footnote{Waismann (1951a), p. 54. See also Waismann (1951b), pp. 115-121.} The truths of logic, by contrast, are not vague in this respect—they are all transparently necessary (i.e., they admit of no alternatives) and \textit{a priori} (i.e., empirical evidence is not relevant to their rational acceptability).

“Language,” Waismann says, “is an instrument that must, as occasion requires, be bent to one’s purposes.”\footnote{Waismann (1952), p. 6.} In many cases, those purposes cannot be achieved except by transgressing the prevailing norms of usage. It is a mistake to think that widely endorsed substitutions will hold in all possible contexts, and we should therefore always be prepared to deviate from ordinary usage as needed. This need is most pressing when we are trying to describe newly discovered phenomena or confronting some otherwise unforeseen circumstance for which our current language is descriptively inadequate. Waismann writes:

\begin{itemize}
\item \footnote{Waismann (1950), p. 25.}
\item \footnote{Waismann (1951a), p. 54. See also Waismann (1951b), pp. 115-121.}
\item \footnote{Waismann (1952), p. 6.}
\end{itemize}
Notice with what unerring instinct language contrives to say, at the cost of a slight departure, what would be unsayable if we moved along the rigid grooves of speech. Indeed, how should one describe such phenomena if not by breaking away from the clichés? Is there anything objectionable in that? If so, language could never keep pace with life. Yet new situations, unforeseen, arise, and with them the need of describing them; it can only be met by adjusting language—either by coining new words, or, as the word-creating faculty is scanty, by pressing old ones into new services, in this way cutting through the dead mass of convention.42

Examples of this sort of creative departure include, inter alia, Einstein’s reconceptualization of simultaneity, Flaubert’s non-conjunctive use of ‘and’, and Hume’s modernization of the term ‘impression’.43 Each of these innovations enriches our shared language, providing us with new means of expression better suited to our scientific, literary, and philosophical needs.

1.5 The Three Philosophical Roles of Analyticity

In light of its history, three essential roles can be discerned for the notion of analyticity. These tasks are interrelated, but importantly distinct. I’ll call the three philosophical roles played by analyticity the metaphysical role, the epistemological role, and the interpretive role, respectively.

In its metaphysical role, analyticity has been called upon to explain the source and necessity of mathematical, logical, and conceptual truths. Carnap and Ayer are the most obvious advocates of this strategy, though Kant and Frege pursue limited versions of it as well. (Waismann is the odd man out here, holding that the necessary/contingent distinction can be drawn clearly only for logic and mathematics, and insisting that it becomes vague and context-dependent when we consider analytic sentences that feature open-textured terms.)

42 Ibid., p. 8.
43 Ibid., pp. 8-13. Waismann also emphasizes the value of syntactical or grammatical innovations and discusses a wide range of examples including Flaubert’s use of tense, Freud’s transitive use of the verb ‘erinnern’, and Lichtenberg’s remarks on the Cogito. See Waismann (1953), pp. 80-81.
Kant thinks that analytic judgments inherit their necessity from the law of contradiction, so it is fair to say that for Kant, a judgment’s status as analytic explains its necessity. Similarly, Frege thinks that since arithmetic can be reduced to logic, the truths of arithmetic are necessary if the truths of logic are. Of course, unlike Carnap, neither Kant nor Frege attempt to use analyticity to explain the necessary status of the logical truths themselves.

In its epistemological role, analyticity has been called upon to explain the *a priori* knowability of mathematical and logical truths, as well as the *a priori* knowledge thought to be gained from conceptual analysis. For Kant, careful examination of the subject-concept in an analytic judgment will reveal that the predicate-concept is already being thought in the subject-concept, and this will result in *a priori* knowledge that the judgment in question is true. For Frege, the discovery of a proof of a proposition which is based solely on general logical laws and definitions will result in *a priori* knowledge that the proposition in question is true. On Carnap’s view, the *a priori* status of mathematics, logic, and any of the various conceptual truths is guaranteed by the conventionally stipulated semantical rules of the language one has adopted.

One obvious connection between analyticity’s metaphysical and epistemological roles that has likely contributed to their being conflated and run together is that certainty, indubitably, and empirical indefeasibility appear guaranteed if the object of belief is a necessary truth. However, in light of the possibility of *a posteriori*, yet necessary, truths—like ‘Water is H₂O’ or ‘Temperature is mean molecular kinetic energy’—most philosophers today see *a priority* and necessity as coming apart in various cases. For the *locus classicus* of this view see Kripke (1980).
these two roles can and should be pulled apart. But there remains further pulling apart to be
done.

While analyticity’s metaphysical role is exemplified in the characterization of an
analytic sentence as one that is “true by virtue of meaning alone,” the epistemological role is
exemplified in the characterization of an analytic sentence as one in which “understanding of
its meaning suffices for knowledge of its truth.” There is, however, a third role played by
analytic sentences—the interpretive role—which is distinct from, though related to, the
epistemological role. The following characterization best exemplifies the interpretive role: an
analytic sentence is one that is meaning-constitutive with respect to one or more of its
component expressions. Thus, putting the three characterizations together, we get this
formula: An analytic sentence is a constitutive principle that is a priori knowable and true by
virtue of meaning alone.

A sentence plays the interpretive role by (a) guiding our interpretive expectations
with respect to the proper use of shared expressions; (b) providing a pro tanto standard for
assessments of linguistic understanding; and (c) providing a pro tanto standard for
distinguishing changes of meaning from changes of belief. An excellent example of the
interpretive role in action is discussed by Grice and Strawson in their seminal reply to
Quine’s attacks on analyticity.\textsuperscript{45} Grice and Strawson ask us to consider someone (“Mr. Y”) who sincerely asserts something that implies the negation of the sentence ‘No three-year-old child is an adult’. That is, Mr. Y asserts ‘My neighbor’s three-year-old child is an adult’, and subsequently refuses to accept paraphrases such as ‘You mean he’s very advanced for his age’ or ‘You mean he won’t grow anymore’, etc. According to Grice and Strawson,

\textsuperscript{45} Grice and Strawson (1956), p. 205.
At this stage—or possibly if we are patient, a little later—we shall be inclined to say that we just don’t understand what Y is saying, and to suspect that he just does not know the meaning of some of the words he is using. For unless he is prepared to admit that he is using words in a figurative or unusual sense, we shall say, not that we don’t believe him, but that his words have no sense. And whatever creature is produced for our inspection, it will not lead us to say that what Y said was literally true, but at most to say that we now see what he meant.\[^{46}\]

Grice and Strawson’s discussion of this case is important because it draws attention to the fact that when we treat a sentence like ‘No three-year-old child is an adult’ as an interpretation-guiding constitutive principle governing the proper use of the word ‘adult’, our acceptance of this standard is manifest in our interpretive behavior. In some contexts, we are disposed to treat those who endorse the negation of that sentence as *prima facie* incompetent with the expression ‘adult’; in other contexts, we take this as evidence of a difference in meaning. In this way, the sentence ‘No three-year-old child is an adult’ plays what I have called the interpretive role.

### 1.6 Is Analyticity a Legitimate Theoretical Notion?

Although the notion of analyticity has traditionally been expected to play each of the three philosophical roles discussed above, there are important reasons for skepticism. The first argument against analyticity that should be considered resists easy classification as a criticism of any one of the three roles in particular. Instead, it purports to be an argument against analyticity’s ability to play *any* role in serious philosophizing. The criticism to which I am referring is Quine’s infamous “circularity objection” from “Two Dogmas of Empiricism.”\[^{47}\]

\[^{46}\] Ibid.
\[^{47}\] Quine (1951). See also White (1950). For early responses, see Mates (1951), Martin (1952), and Kaufman (1953).
The argument goes something like this. If a notion is to do important philosophical work, it must be possible to give a sufficiently clear, non-circular definition of it—that is, it must be wholly definable in terms of other concepts on which we have a firm cognitive grip. Analyticity, however, cannot be so defined—any attempt to define it will inevitably appeal to some other, equally murky intensional notion, such as necessity, or synonymy, or semantical rule, “which is no less in need of clarification than analyticity itself.” That is, in saying what analyticity is, we cannot help but appeal to at least some of these other notions, and in spelling them out, we cannot avoid appealing ultimately to analyticity. Therefore, according to this line, analyticity is not eligible for use in serious philosophical theorizing.

One might argue that Quine’s circularity objection is unconvincing on the grounds that it depends upon a dubious presupposition: viz., the idea that only notions which can be specified in austere, physicalist or behaviorist terms are suitable for serious theorizing. If this is right, Quine’s so-called circularity argument would seem to beg the question against anyone who isn’t already a die-hard physicalist or behaviorist. It is not clear, however, that this is the most charitable way to interpret the objection. It is true that, in light of his naturalistic commitment to methodological behaviorism with respect to linguistic theory, Quine himself would reject certain ways of trying to break out of the circle which would be regarded as legitimate by non-Quineans. But the objection is not itself predicated on this commitment. Instead, Quine’s real worry seems to be that the notions in the intensional family to which analyticity belongs—including meaning, definition, synonymy, necessity, essence, and semantical rule—are not genuinely explanatory notions at all. Quine’s point is that any notion which is a member of a family that exhibits the sort of inter-definitional circularity

---

48 Quine (1951), p. 23.
characteristic of the analyticity group cannot be supposed to play an essential explanatory role in an account human linguistic practice unless there is some way to break out of the circle. The only way to do this would be to reduce one of the notions in the family, but in light of the tight inferential connections that bind the family together, no proposed reduction will appear plausible.

In their reply to Quine, Grice and Strawson note two other examples of circular families: what we might call the morality group (morally wrong, morally right, blameworthy, moral rule); and what we might call the assertion group (true, false, statement, fact, denial, assertion). Surely, they argue, the notions that make up the morality group and the assertion group are not senseless or unintelligible just because they are inter-definable and resist reduction. Grice and Strawson are clearly right that such expressions are perfectly meaningful and intelligible. However, questions of meaningfulness and intelligibility are beside the point here. The real question is whether or not moral, modal, or semantic vocabulary is suitable for the purposes of explanatory theorizing.

1.7 Can Conventional Stipulation Explain Necessary Truth?

In “Truth by Convention,” Quine undermines analyticity’s ability to play the metaphysical role by arguing that conventional definition or stipulation cannot even explain the truth of a given sentence, let alone its necessity. The core idea here is that no sentence can be made true simply by definitional fiat—if a sentence is true, it is true because of the way the world is, not because of any stipulative activity on our part. We can, of course, offer up


\[50\] See chapter 2 of Quine (1960) for his attempt to show that the (unreduced) notions in the analyticity group are not needed in order to explain what goes on when human interpreters translate one another’s speech.

\[51\] Quine (1936).
stipulative definitions—for example, ‘⌘p = df ¬(p ∧ ¬p)’—from which various truths will follow, for example, ‘⌘p ↔ ¬p ∨ ¬¬p’—but we have not thereby created any new source of truth or provided anything like an explanation of necessity. All that stipulative definition can do is introduce an abbreviation or alternative notation for some existing true sentence; as such, whatever explains the truth or necessity of the definiens also explains the truth or necessity of the definiendum.52

Quine reminds us that “true sentences generally depend for their truth on the traits of their language in addition to the traits of their subject matter”53 and he argues that there is no reason to think that purportedly analytic sentences are any different in this respect.

Gilbert Harman puts the point nicely:

According to the full blooded theory of analytic truth, there are... sentences that express truths solely by virtue of their meanings and independently of the way the world is. The sentences “Copper is a metal” and “Copper is copper” would be said to express truths solely by virtue of their meanings. ... The problem with this lies in understanding how the truth of a sentence can be independent of the way the world is and depend entirely on the meaning of the sentence. What is to prevent us from saying that it is a fact about the world that copper is a metal and that, if this were not so, the sentence “Copper is a metal” would not express a truth? And what is to prevent us from saying that the truth expressed by “Copper is copper” depends in part on a general feature of the way the world is, namely that everything is self-identical?54

The advocate of analyticity is thus challenged to defend the explanatory claim that some sentences are true by virtue of meaning alone, irrespective of how the world is. In order to answer Quine and Harman, the defender of analyticity must do two things. First, she must cash out precisely what is meant by the claim that the truth of analytic sentences is independent of worldly facts. That is, she must show that the analyticities constitute a class of sentences

52 Ibid, pp. 78-88.
53 Quine (1963), p. 113.
that are true in a special way, making them fundamentally different in kind from ordinary synthetic truths. Second, she must show that her explanatory hypothesis is superior to the Quine-Harman alternative according to which the truth-value of any sentence always depends in part upon the way the world is.

Quine’s attack on logical conventionalism severely undermines the plausibility of the idea that analyticity can play the metaphysical role. For if stipulation creates no new truths but merely introduces abbreviations of old truths, then what is metaphysically special about the class of so-called analytic statements? The answer cannot be that they are stipulated to be true by convention, since this does not explain their truth and thus does not distinguish them from synthetic statements in any significant way. Unless we cleave to the (quite implausible) view that stipulative activity actually generates a special class of truths—that certain statements are made true by the adoption of certain conventions—it is hard to see how the analytic-synthetic distinction could play anything like the metaphysical role traditionally expected of it.

1.8 Can Implicit Definition Ground A Priori Knowledge?

Suppose, following Carnap, that we decide to adopt a set of semantical rules which we take to define or constitute a certain linguistic framework. Let us grant, moreover, that by enumerating this list of statements and stipulating them to be analytic, we have thereby explained their truth and necessary modal status along conventionalist lines. Even granting this much, however, the task of explaining how such statements are knowable a priori still remains.
Carnap holds that the logical constants acquire their meanings through the process of implicit definition. First, we introduce a new sub-sentential expression, say, the operator ‘→’. Next, we stipulate that ‘→’ is to mean whatever it needs to mean in order for a certain set of sentences Γ to be true. The set Γ will include the introduction and elimination rules for ‘→’, and various other logical truths in which ‘→’ appears. The meaning of ‘→’ is thus constrained by the requirement that all the members of Γ come out true. Moreover, since we know of the meaning of ‘→’, we are now in a position to know certain logical truths a priori—e.g., statements of the introduction and elimination rules for ‘→’—because of their role as implicit definers. This process can then be applied to the rest of the logical constants in our language, yielding similar results.

Quine objects to the implicit definition strategy by arguing that any attempt to use this mechanism to account for the a priori of logical truths in general is bound to illegitimately assume the truth of at least some logical principles. Consider the case of ‘→’ described above. Since there will be an infinite number of logical truths that are members of Γ, the explanation of how ‘→’ is implicitly defined will have to appeal to schemata. Among the schemata that need to be true in order for ‘→’ to acquire its meaning is ‘((p → q) ∧ p) → q’. Let’s call this schema S. The reasoning would then have to go something like this:

1. For any sentence x, if x is an instance of schema S, then x must be true
2. Sentence ϕ is an instance of schema S

(3) ϕ is true

---

55 I am glossing over the distinction between inference rules and logical truths. For my purposes here, it should be enough to note that inference rules can be put into the form of truth-functional sentences called transfer principles. The following is a transfer principle for Modus Ponens: ‘((p → q) ∧ p) → q’. Any transfer principle derived from a valid inference rule will turn out to be a logical truth.

56 See Quine (1936) p. 103.
Notice, however, that in order for this inference to be any good, we have to assume the validity of an inference rule—e.g., universal instantiation—which we have not yet proved. This general situation apparently cannot be avoided, since any attempt to formulate a line of reasoning that would validate a given inference rule must assume the legitimacy of inferential transitions that have yet to be validated.

It is not exactly clear that the circularity involved here is truly pernicious. Perhaps this kind of “rule-circularity” is only problematic when one antecedently doubts the inference rule in question. Plausibly, whenever one has no reason to doubt the joint validity of a pair of harmonious introduction and elimination rules featuring the same concepts, one is entitled to infer in accordance with them. So, we may still be warranted in reasoning in accordance with these basic inference patterns even though positive justification in any stronger sense can never be provided.

Nevertheless, the crucial point is that the implicit definition strategy cannot account for the particularly robust kind warrant that our most basic logical inference patterns appear to enjoy. Defeasible entitlement to provisionally assume the truth of allegedly analytic sentences in a particular context is not the same as genuine a priori warrant for believing genuinely analytic truths. Harman provides clarity on the issue:

In a particular inquiry certain premises may be taken for granted and not questioned. We could say such premises are known a priori, i.e., at the beginning of inquiry, while other things, discovered in the course of the inquiry, come to be known only a posteriori. But this would not mean that we have a priori knowledge of the premises of the inquiry in any sense usable by the philosophical defender of analyticity. For these premises need not be known solely by virtue of knowledge of their meaning. They may well be known as the result of prior empirical inquiry. They may not be known at all, but only assumed to be known. The defender of analyticity needs more than such relatively a priori knowledge. He needs absolutely a priori knowledge, whatever that would be.

Arguably, an essential feature of “absolutely” a priori knowledge is empirical indefeasibility. For unless the warrant generated by grasping the meaning of analytic sentences is empirically indefeasible, analyticity cannot provide a satisfactory ground for our knowledge of logical, mathematical, or conceptual truths—truths which are supposed to hold no matter what the world is like. Unfortunately, there are countless examples of sentences which have been taken to be paradigm cases of analytic truths—for example ‘Atoms have no parts’ and ‘Any two events that occur at the same time happen simultaneously’—which have turned out, in light of subsequent empirical inquiry, to be false. Of course, the defender of analyticity can say, in retrospect, that such sentences must not have been analytic after all. But note that to say this is to admit that humans are wildly unreliable when it comes to detecting genuine analyticity. And if we aren’t reliable detectors of genuine analyticity, then how can we be sure that we have genuine a priori warrant in any particular case?

Appeals to implicit definition thus fail to establish that logical or conceptual truths are knowable with a sufficiently robust sort of warrant to qualify as genuinely a priori. Our entitlement to hold true purportedly analytic sentences has the same epistemic force as our entitlement to make assumptions for the sake of argument—in either case, it is simply a matter of what our conversational partners will let us get away with, modulo our respective sets of background beliefs and shared conversational goals. This weak, empirically defeasible, conversation-relative form of warrant is a far cry from the robust, absolute, empirically indefeasible variety of a priori justification that analytic sentences would have to provide in order to successfully play the epistemological role.
1.9 Can Analyticity Explain Linguistic Understanding?

It is natural to think that in order to count as competent with a certain expression, one must believe or accept as true the analytic principles one takes to specify the core meaning of that expression. However, there are some cases in which the putatively analytic sentences that serve to constitute the meaning of an expression are jointly inconsistent.\(^{59}\)

Perhaps the most salient example of this is the case of the notion of truth. Two principles govern our concept of truth: “From ‘p’, infer ‘p is true’”, and “From ‘p is true’ infer ‘p’”. But although they would seem to be analytic, these principles cannot both be true (and thus cannot be known \textit{a priori}). As Kevin Scharp has argued, the incompatibility of truth’s constitutive principles is precisely what generates the liar paradox, yet our competence with the concept \textit{truth} puts pressure on us to accept both of these principles.\(^{60}\)

To take a different example, consider the following case from Anil Gupta. Imagine a community with a certain concept—i.e., \textit{up above}—which is governed by two criteria: a perceptual criterion (which authorizes the assertion of ‘\(a\) is up above \(b\)’ in obvious perceptual situations), and a conceptual criterion (which authorizes the assertion of ‘\(a\) is up above \(b\)’ just in case the direction of the ray \(ba\) is the same direction as Standard Up).\(^{61}\) Because of the spherical shape of the earth, it turns out that “[f]or many pairs of objects \(a\) and \(b\), the rules dictate that the assertion of ‘\(a\) is up above \(b\)’ is warranted and also that its denial is warranted.”\(^{62}\) Thus, it is not the case that one’s understanding of the expression ‘\(\text{up above}\)’ is sufficient for \textit{a priori} knowledge of the principles that govern that expression. However, the

\(^{59}\) See Eklund (2002).


\(^{61}\) Gupta (1999). Standard Up is defined as the direction of the ray \(ps\), where \(p\) is a certain salient mountain peak and \(s\) is a certain salient satellite which happens to be in geostationary orbit above \(p\).

\(^{62}\) Ibid., p. 20.
fact that one remains competent with the expression even after becoming aware of the conflict in the principles entails that linguistic understanding comes apart from a priori knowledge.

Another source of discomfort about the alleged connection between believing analytic truths and understanding linguistic expressions comes from reflection on two thought-experiments proposed by Tyler Burge. In the first thought-experiment, Burge shows that a person may count as (partially) competent with a public expression like ‘arthritis’ even if that person has some false beliefs about where arthritis can occur and thus does not believe the apparently analytic sentence ‘Arthritis is a disease of the joints only’. Burge argues that so long as the person in question is disposed to defer to the relevant experts when his usage is challenged, he counts as someone who understands of the word ‘arthritis’, albeit incompletely.63

In addition to such cases of incomplete understanding, Burge also argues that a person might understand a term perfectly well, yet endorse a radically non-standard theory about the thing or things to which that term refers.64 In this second thought-experiment, Burge suggests that a person could conceivably have the ability to understand all sorts of claims about sofas—and thus be fully competent with the term ‘sofa’—and yet, on the basis of her acceptance of a non-standard theory according to which sofas are delicate religious objects devoid of practical functionality, refuse to believe the apparently analytic sentence ‘Sofas are pieces of furniture made or meant for sitting’.65 Thus, according to Burge,

---

64 Burge (1986).
65 Ibid., pp. 707-709.
“necessary truths that intuitively give the meaning of an empirically applicable term (or purport to provide a general understanding of such a concept) are dubitable.”

More recently, Timothy Williamson has formulated several examples which are very similar to Burge’s cases. He imagines two deviant logicians—Peter and Stephen—who obviously possess the concept *vixen*, but who refuse to accept the putatively analytic, intuitively meaning-constitutive sentence ‘Every vixen is a female fox’. Williamson describes Peter’s views as follows:

[Peter] regards the truth of ‘There is at least one F’ as a necessary condition for the truth of ‘Every F is a G’ quite generally, and the falsity of ‘There is at least one F’ as a sufficient condition for the falsity of ‘Every F is a G’; he takes universal quantification to be existentially committing. … Peter also has the weird belief that ‘There is at least one vixen’ is false. For he spends far too much time surfing the internet, and once came across a site devoted to propagating the view that there are no foxes, and therefore no vixens, and never have been: all the apparent evidence to the contrary has been planted by a secret international agency; for sinister purposes best known only to itself, it produces elaborate fox-hallucinations. Being a sucker for conspiracy theories, Peter accepted this one.

He then describes Stephen’s views:

Stephen has no time for Peter’s pet theories. What worries him is vagueness. He believes that borderline cases for vague terms constitute truth-value gaps. … On Stephen’s view, for ‘Every F is a G’ to be true is for the conditional ‘x is an F → x is a G’ to be true for every value of the variable ‘x’; for ‘Every F is a G’ to be false is for conditional ‘x is an F → x is a G’ to be false for some value of ‘x’. … Stephen also believes that some clearly female evolutionary ancestors of foxes are borderline cases for ‘fox’ and therefore ‘vixen’. Consequently, for such an animal as the value of ‘x’, ‘x is a vixen’ is neither true nor false, so the conditional ‘x is a vixen → x is a vixen’ is also neither true nor false, by the strong Kleene table for →.

These cases further support the contention that one can understand a certain linguistic expression and yet deny one or more of the putatively analytic principles that govern the use of that expression.

---

66 Ibid., p. 699.
69 Ibid., pp. 10-11.
Intuitively, ‘Every vixen is a female fox’ is an analytic truth which gives meaning to the term ‘vixen’—that is, the sentence seems to specify the very essence of vixenhood, or what it is to be a vixen. As such, to deny that all vixens are female foxes would seem to betray a lack of competence with the term ‘vixen’ or a failure to grasp what is essential to being a vixen. Yet, as we consider Williamson’s description of his characters’ reasons for denying what appears to be a conceptual truth, it does not seem right to say that Peter or Stephen lack competence with terms like ‘vixen’ or ‘female fox’. Even if they lack the relevant dispositions themselves, Peter and Stephen are certainly well aware of what someone’s dispositions to apply ‘vixen’ would have to be like in order to count as competent with that expression.

It is less obvious, however, that Peter and Stephen are entitled to claim that they are using the quantificational idioms ‘all’ and ‘every’ with the same meaning as ordinary competent speakers of English. Williamson, of course, insists that his characters are not merely changing the subject. Instead, they are recommending that we change our usage so that it better reflects the true nature of its subject matter. We are meant to see them as fully competent speakers of a shared public language who hold incompatible substantive theories about the nature of quantification. One apparent consequence of seeing things this way, Williamson suggests, is that we must give up on the idea that there is any useful distinction to be drawn between meaning-constitutive and non-meaning-constitutive sentences.

70 Putnam (1962a) calls bachelor and vixen “one-criterion” concepts because their meaning appears to be exhaustively characterized by a single conceptual truth. By contrast, he calls scientific concepts like mass or kinetic energy “law-cluster” concepts, since their meaning cannot be exhaustively characterized with a single criterion, in light of the theoretical role such concepts play in science. Williamson’s choice of examples is designed to undermine the idea that there are conceptual truths even in the intuitive case of so-called “one-criterion” concepts.

71 See McGee (1985) for a real-life example of an expert who holds a non-standard theory about conditionals.
1.10 Concluding Remarks

In my view, the examples from Burge and Williamson discussed in the previous section do not conclusively show that nothing is meaning-constitutive. Such cases are instructive, however, for they demonstrate the need for a dynamic, rather than a static, account of what counts as a constitutive principle. As Davidson has stressed, competence with a linguistic expression should not be understood in terms of a particular set of core beliefs, assent to which is necessary and sufficient for understanding the expression in question. For although there must be sufficient overlap between what the interpreter believes and what her interlocutor believes, there is always room for negotiation as to exactly which sentences or inferences featuring expression \( E \) are (or are not) to be counted as constitutive of the meaning of \( E \). With sufficient specification of context, we can grant that Williamson’s characters are competent English speakers, yet consistently maintain that the relevant disputes with them are best explained as disputes over which principles are properly treated as meaning-constitutive with respect to quantificational terms like ‘all’ and ‘every’. Depending upon how that context is specified, it may or may not be appropriate to think of Peter or Stephen as proposing a change of subject rather than advocating incompatible substantive theories.

What, then, of the interpretive role allegedly played by analytic sentences? Considerations about inconsistent concepts, incomplete competence, and competent-yet-deviant theorists do not entail that no sentence actually plays the interpretive role. What such considerations do suggest is that (a) sentences which serve as interpretation-guiding

---

principles need not be true; and (b) competent speakers need to be able to discern which sentences their peers treat as interpretation-guiding, but need not actually believe them. The upshot of this is that the legitimacy of the practice of sorting constitutive from non-constitutive sentences does not appear to depend on the theoretical viability of the analytic-synthetic distinction. It is thus possible to agree with each of the four major criticisms of analyticity discussed above and still insist that distinguishing between constitutive and non-constitutive uses of our expressions is an essential aspect of competent linguistic practice.
Chapter 2

Three Recent Accounts of Analyticity: A Critical Survey

2.1 Introduction

Keeping in mind the important criticisms of analyticity discussed in the previous chapter, I now turn to a consideration of three contemporary attempts to resurrect some version of the notion. First, I examine Paul Boghossian’s account of “epistemic analyticity,” then I consider Gillian Russell’s account of “truth in virtue of meaning,” and finally, I assess the prospects of “analyticity*,” the replacement notion proposed by Cory Juhl and Eric Loomis. It will be argued that none of these three recent proposals is successful. Although each account is problematic in specific ways, they share an important general failing: viz., they each fail to deliver a notion that is capable of doing the sort of philosophical work that it is analyticity’s job to perform.

2.2 Boghossian on “Epistemic Analyticity”

Paul Boghossian defends an account of analyticity which places primary emphasis on analyticity’s epistemological role while rejecting the idea that analytic sentences are true in virtue of meaning alone. According to Boghossian, the idea that analyticity can explain the source of necessary truth has been decisively demolished by Quine, but a revised notion—
“epistemic analyticity”—can, he thinks, survive Quinean attacks. Boghossian’s overarching goal is to put his notion of analyticity to work in an account of a priori knowledge, including knowledge of logic.

Analytic sentences are supposed to be epistemically special in that simply understanding such a sentence is sufficient for a priori knowledge of its truth. If all analytic sentences have this special epistemic feature, then perhaps analyticity is the key to explaining how a priori knowledge in general is possible. But in order to give an account of how analyticity might perform such important work, one will need to first give an account of how understanding of meaning is supposed to produce a priori justification. A classic way of doing this is to invoke, as Boghossian does, the idea of implicit definition.

On Boghossian’s account, epistemic analyticity comes in two varieties, which he labels Frege-analyticity and Carnap-analyticity. A sentence counts as Frege-analytic just in case it is transformable into a logical truth by the substitution of synonyms for synonyms.73 Hence, a sentence like ‘Vixens are female foxes’, is Frege-analytic because, given the synonymy of the expressions ‘vixen’ and ‘female fox’, the sentence can be transformed into a logical truth, viz., ‘Female foxes are female foxes’. As a result of this transformability, Frege-analytic sentences will be a priori knowable provided that the logical truths into which they are transformable are a priori knowable. Although Frege-analyticity cannot by itself account for the a priori status of logical truths, Boghossian thinks that the other variety of epistemic analyticity—viz., Carnap-analyticity—can. A sentence counts as Carnap-analytic just in case it is an implicit definer with respect to one or more of the expressions that it contains.74 To

73 Boghossian (1997), p. 337. It should be noted that this label is misleading, since Frege himself does not characterize analyticity in terms of synonymy.
74 Ibid., p. 339.
say that a sentence (or inference) implicitly defines an expression it contains is to stipulate that the sentence (or inference) must be true (or valid) if the expression in question is to have the meaning that it does. For example, to say that the standard introduction and elimination rules for conjunction implicitly define the expression ‘&’ is to say that ‘&’ means whatever it has to mean in order for those introduction and elimination rules to come out valid. Because logical truths are plausibly thought of as implicit definers of the logical expressions they contain, the logical truths will count as Carnap-analytic.

The key task for Boghossian, then, is to demonstrate how implicit definition can explain our a priori knowledge of logic. The standard way to do this is to combine the implicit definition account with the doctrine of conventionalism—i.e., the idea that by making the stipulations required for the implicit definition of our logical expressions, we thereby generate a special class of conventional truths which we can confidently claim to know a priori precisely because our stipulative intentions are responsible for their very existence. Since Boghossian rejects analyticity’s metaphysical role, he cannot avail himself of the conventionalist strategy. He can, however, consistently endorse the doctrine of implicit definition without also endorsing conventionalism. For although the theory of implicit definition was originally accompanied by a commitment to conventionalism, Boghossian argues that the two views are independent of one another: the fact that meaning is fixed by implicit definition does not imply that the implicitly definitional sentences and inferences are true by convention. About this, Boghossian seems to be quite right; for it might well be the case that our taking a certain set of sentences featuring expression $E$ to be true is what constitutes $E$’s meaning, yet this in no way implies that those sentences somehow become true.
as a consequence of our taking them to be so. An implicit definer, like any other sentence, is true just in case the world is as it says it is.⁷⁵

Although Boghossian can consistently hold on to implicit definition without also being forced to accept conventionalism, he must still provide an explanation of how we get from implicit definition to a priori knowledge of logic. Indeed, this task becomes especially urgent in light of his wholesale rejection of analyticity’s metaphysical role. In general, Boghossian seems disposed to reject any explanatory strategy according to which a priori knowledge is possible because its subject matter is mind-dependent or constituted by intentional human activity. Both Kant’s treatment of the a priori and the positivists’ conventionalism are instances of this general strategy. In an attempt to navigate between the Scylla of rationalist excess and the Charybdis of empiricist austerity about the a priori, Kant ingeniously proposes that the truths (both analytic and synthetic) to which we have a priori access are knowable in this distinctive way because of a certain kind of mind-dependence. Analytic truths, for Kant, are truths about containment relations between our concepts, and since we have introspective access to the contents of our concepts, we can know these containment relations a priori.⁷⁶ Synthetic a priori truths, while not purely conceptual, are still about something that is contributed by the mind—viz., our pure intuitions of space and time—and this mind-dependence explains how we can have direct access to them and confidently recognize their necessity.⁷⁷

⁷⁶ See Critique of Pure Reason (A6-7/B10-11).
⁷⁷ Ibid., (B73).
Conventionalism is an attempt to address the same basic Kantian problematic about the *a priori*—only without the resources afforded by transcendental idealism.\(^{78}\) Recall that on Boghossian’s understanding of conventionalism, our stipulative intentions generate the truths of meaning—special truths that are metaphysically dependent upon and constituted by human activity. Just as Kant’s view is designed to explain how human beings can have reliable epistemic access to a special class of (mind-dependent) truths, conventionalism is designed to explain how human beings can have reliable epistemic access to a special class of (conventional) truths. The idea is that we set up conventions by intending to use expressions in certain ways, and if we transparently know what those intentions are, then we have direct epistemic access to the source from which analytic truth springs. On such a view, analytic truths *come to be* as a result of our linguistic intentions, and the fact that we are responsible for their very existence makes them excellent candidates for *a priori* knowability.

So how does Boghossian propose to account for our supposed special epistemic access to analytic truths if he cannot appeal to conventionalism (or transcendental idealism)? His answer is that we are entitled to infer in accordance with implicitly defining inference rules just in case the concept to which those rules give meaning is *epistemically non-defective*. Examples of entitlement-precluding defective concepts include Arthur Prior’s infamous connective ‘tonk’ and the racial pejorative ‘boche’ discussed by Michael Dummett.\(^{79}\) The (purported) connective ‘tonk’ is problematic because its introduction rule permits the

---

\(^{78}\) There are, of course, important differences, but they are differences of detail. Although he gives an account of both the analytic and the synthetic *a priori*, Kant was most interested in the question of how there could be synthetic *a priori* judgments. The positivists rejected the notion of the synthetic *a priori* and the transcendental idealism required to accommodate it. But they still thought that the analytic variety of *a priori* knowledge was possible, and sought to provide an explanation of it grounded in human activity. Viewed in this way, conventionalism is a post-Kantian strategy proposed by philosophers working squarely in the Kantian tradition. For more on the historical ties between Kant and logical positivism, see Coffa (1991).

\(^{79}\) For ‘tonk’, see Prior (1960); and for ‘boche’, see Dummett (1973), p. 454.
inference from ‘A’ to ‘A tonk B’ and its elimination rule permits the inference from ‘A tonk B’ to ‘B’. Thus, inferring in accordance with the rules for ‘tonk’ allows one to derive any arbitrary premise from any other. The racial pejorative ‘boche’ has a similar (if less severe) problem: it permits inferences from ‘x is German’ to ‘x is boche’; and from ‘x is boche’ to ‘x is cruel’. Thus, inferring in accordance with the rules for ‘boche’ allows one to conclude of any arbitrarily chosen German that he or she is cruel, thereby committing the user of ‘boche’ to the (empirically false) generalization that all Germans are cruel.

Boghossian notes that while we might plausibly deny that ‘tonk’ successfully denotes a genuine concept, “it’s hard to believe that racists who employ boche-like concepts fail to express complete thoughts.” 80 In fact, he suggests that the problem with boche is not simply that it is governed by rules that fail to be truth-preserving, but that it requires its users to unconditionally believe a controversial theory about Germans. On the basis of such considerations, Boghossian concludes that, in general, no concept should “be designed in such a way that only those who believe a certain creed are allowed to possess it.” 81 Concepts like boche are epistemically defective because their possession conditions preclude us from raising reasonable doubts about their extensions. To be non-defective, these concepts need to be governed by what Boghossian calls conditionalized rules: “rules that conditionalize on the existence of an appropriate semantic value that would make its rules truth-preserving.” 82 Thus, a non-defective version of the concept boche would only commit its users to a conditionalized version of the theory implicit in its introduction and elimination rules: If anything is boche, then all Germans are cruel. The idea is that the use of this conditionalized

80 Boghossian (2003), p. 28.
81 Ibid., p. 29.
82 Ibid., p. 31.
version of the concept would not directly commit its users to an empirically false theory about Germans because they would be free to deny the antecedent of the conditionalized statement of the theory.

However, there will be some concepts, according to Boghossian, which are necessarily unconditionalized, but nevertheless epistemically non-defective. These concepts count as non-defective, according to Boghossian, because they “will be presupposed in any conditionalization and those that are so presupposed will not themselves have conditionalized versions.”

Boghossian’s reasoning seems to be something like this. Because we are rational, we (ought to) want to ensure that the concepts we use are epistemically non-defective; for we are only entitled to employ concepts that are non-defective. One way to ensure that a concept is non-defective is to make sure that we use a conditionalized version of it. But in order for us to have any conditionalized concepts at all, we need to have already grasped certain basic logical concepts (i.e., those concepts one must grasp in order to understand the conditionalization process itself). For this same reason, these basic logical concepts are only available to us in unconditionalized versions. So, in order to ensure our entitlement to employ all of our other concepts, we must reason in ways that presuppose the legitimacy of certain unconditionalized basic logical concepts. Therefore, there are at least some unconditionalized concepts that are epistemically non-defective, namely, the basic logical concepts that one must grasp in order to understand concept-conditionalization itself.

Boghossian’s big move here boils down to the claim that because we cannot help but employ unconditionalized basic logical concepts, we are thereby entitled to do so. On this account, we are entitled to infer in accordance with the standard introduction and

---

83 Ibid., p. 32.
elimination rules for the standard connectives because if we weren’t, we wouldn’t be able to
formulate justifications of anything or offer reasons for any of our beliefs. But although this
entitlement does not depend on empirical evidence, Boghossian must admit that it does not
provide a warrant that is empirically indefeasible. This is problematic for his account, since
empirical indefeasibility is a classic hallmark of the *a priori*. Perhaps Kant is right that our
minds are structured such that we cannot help but experience space as having exactly three
dimensions. If so, perhaps we are thereby entitled to believe that space has exactly three
dimensions. But for all that, the claim that space has exactly three dimensions is certainly not
empirically indefeasible, and it does not seem to be something we can know *a priori*.

In his earlier discussions analyticity, Boghossian distinguishes between a “minimal
notion” of *a priori* warrant (according to which a belief counts as *a priori* if it is justified
“without appeal to empirical evidence”) and a “stronger notion” (according to which “the
justification in question is not defeasible by any future empirical evidence”). He claims that,
in his view, “the minimal notion forms the core of the idea of apriority”, but goes on to say
that “under the appropriate circumstances, the notion of [epistemic] analyticity can help
explain how we might have a priori knowledge even in the strong sense.”84 Despite such
promises, Boghossian’s account can at best establish that we have “minimal” *a priori* warrant
for our logical beliefs, in the sense that our entitlement to logic does not directly appeal to
empirical evidence. Nothing in his account implies that there are ever circumstances in
which we have “strong” *a priori* warrant for our logical beliefs. Although Boghossian claims
that this minimal notion of *a priori* warrant forms the “core” of the idea of *a priority*, his
abandonment of the empirical indefeasibility criterion marks a drastic break with tradition.

---

Both Kant and the logical positivists would have been deeply unsatisfied with any account of the *a priori* of logic that failed to establish its empirical indefeasibility. For many philosophers, past and present, such an account would arguably not count as an explanation of the *a priori* of logic at all.

What Boghossian provides us, it seems to me, is not an account of the *a priori* of logic as traditionally conceived, but rather a proposal for conceptual revision. Because the empirical indefeasibility of logic cannot be established, Boghossian urges us to drop that criterion from our notion of *a priori* warrant and operate with a revised concept according to which a warrant counts as *a priori* just in case it does not directly depend on empirical evidence. Arguably, to accept this proposal is simply to give up on the traditional epistemological project for which the notion of analyticity was invoked by Kant and his positivist descendants—it is, in other words, to side with Quine and give up on analyticity's epistemological role. Because Boghossian’s “epistemic analyticity” cannot do the explanatory work for which analyticity was traditionally invoked, Boghossian’s account arguably provides us with nothing more than an ersatz notion of little use or interest.

### 2.3 Russell on “Truth in Virtue of Meaning”

By Gillian Russell’s lights, what matters most for our classification of a sentence as analytic is that it retains the same truth-value across certain kinds of contextual variation, and that it exhibits this truth-value invariance in virtue of (some aspect of) its meaning. On Russell’s view, the phrase “truth in virtue of meaning” is to be understood as truth in virtue of *reference determiner*. Along with *content* and *character*, the *reference determiner* is claimed to be a third aspect of meaning associated with referring expressions. According to Russell, a
reference determiner is “a condition which an object must meet in order to be the referent of, or fall in the extension of, an expression.” Roughly speaking, analytic sentences are those that contain expressions whose reference determiner condition is always met, no matter what the world is like. We might think of a reference determiner for expression E in terms of the following conditional: If object X meets conditions C, then E refers to X. If the antecedent of such a conditional is true in all possible worlds, then some of the sentences featuring E will turn out to be analytic in Russell’s sense. Thus, the sentence ‘All bachelors are male’ comes out analytic (a traditional result), as does ‘I am here now’ (a not-so-traditional result). In the former case, the idea is that no matter what possible world you consider, the objects in the extension of 'bachelor' always meet the condition of being male. In the latter case, the idea is that whenever the indexical terms ‘I’, ‘here’, and ‘now’, are uttered, they always refer to the utterer, the time of utterance, and the location of utterance, respectively.

According to Russell, we establish the rules of reference determination via baptismal ceremonies in which we stipulate that certain words are to refer to certain objects or kinds. The reference determiner is thus “the condition specified by the baptizer (using a description, or by pointing) and used to pick out a referent for the name when it was introduced.” Given the role that stipulation plays in instituting reference determiners on her account, one would expect Russell to spend a good amount of time discussing it. However, she simply takes the notion of stipulation for granted, offering very little discussion of the baptismal process or how it is supposed to work.

---

86 Ibid., p. 47.
87 This criticism is raised by Juhl and Loomis (2010), pp. 240-245.
On Russell’s view, it is possible to be competent with a certain expression, yet not know its reference determiner. As such, linguistic competence does not require being present at (or even being aware of the existence of) the relevant baptismal ceremonies. Moreover, because facts about reference determination are instituted via concrete stipulative acts, knowledge of such facts will be \textit{a posteriori}, not \textit{a priori} (since even the stipulator needs empirical evidence in order to know that the ceremony has occurred). Russell holds that analytic sentences are entailed by reference determiners in conjunction with a basic disquotation principle (i.e., ‘Predicate “F(x)” applies to all and only Fs’). So, \textit{if} you know the reference determiners and the disquotation principle, \textit{then} you are in a position to reason to the conclusion that certain sentences are analytic truths without further empirical observation. But since knowledge of reference determination is not \textit{a priori}, it is not empirically indefeasible, and thus any purportedly analytic sentences entailed by such premises will not be incorrigibly known.

It is most natural to think of Russell’s account as a proposal for \textit{revision or replacement} of the traditional notion of analyticity. In effect, she urges us to preserve the centrality of the metaphysical role, while eschewing both the epistemological and interpretive roles. Russell recognizes this, and claims that “it is only because philosophers conflated character and reference determiner that they thought that sentences that were true in virtue of meaning would be sentences that would grant knowledge to anyone who understood them.”

Moreover, many of the sentences that Russell’s account treats as analytic would not traditionally count as analytic, and this helps contribute to the sense that she has more or less changed the subject. For example, on the assumption that the names it contains successfully

\begin{flushright}
89 Ibid., pp. 66-68.
\end{flushright}
refer, the sentence ‘Mohammad Ali is Cassius Clay’ is supposed to turn out to be analytic for Russell, even though it is neither knowable \textit{a priori} and nor intuitively meaning-constitutive.

But even if what Russell gives us is not what the traditional advocate of analyticity was looking for, her account may still be of interest insofar as it promises to make coherent sense of a philosophically viable notion of “truth in virtue of meaning.” To do this, Russell acknowledges that she has to confront a dilemma.\textsuperscript{90} On the one hand, if she concedes that the truth of an analytic sentence is determined in part by what it means and in part by the way the world is, then the sense in which analytic sentences are true by virtue of meaning becomes trivial, and the primary motivation for drawing the analytic-synthetic distinction in the first place evaporates. For if the truth of analytic sentences is determined in this way—i.e., in the same way that the truth of synthetic sentences is determined—then so-called analytic truths do not constitute a special class of truths at all, and thus the alleged metaphysical distinction between them collapses. On the other hand, if she holds that the world plays no role at all in determining the truth of analytic sentences, then the view entails, absurdly, that analytic sentences aren’t really about the world (or worse, that they are made true by convention).

Russell attempts to avoid this dilemma by arguing that there must be some third sense—a sense that is both properly explanatory and also non-trivial—in which analytic sentences are true by virtue of meaning alone. She tries to motivate this move by asking us to consider a sentence that she (quite idiosyncratically) takes to be a paradigm instance of analyticity:

We can get an intuitive grip on the idea that there is a third option by considering the sentence \textit{I am here now}. Intuitively, the meanings of the expressions in \textit{I am here now} guarantee

\textsuperscript{90} Ibid., pp. 31-32.
that the sentence is true in a way in which the meanings of the words in *snow is white* do not guarantee that *snow is white* is true. This is so even though the meaning of *I am here now* did not make it the case that I am here now. (I'm slaving away at the kitchen table this rainy Monday evening because I skipped out on post-talk drinking to finish my book, not because of the meaning of the English word *I*.) So, intuitively, there is a third, non-trivial (since it doesn't apply to all sentences), yet non-absurd (since it applies to at least one sentence), sense in which the truth of a sentence can be determined by its meaning.\footnote{ Ibid., p. 32, note omitted.}

Russell's idea here is that although worldly facts, many of them contingent, are the real truth-makers for analytic sentences, their meanings also provide an additional guarantee of their truth. But just what kind of guarantee is this? And why think that meaning is providing the relevant guarantee here? Why not say that, given the meaning, the relevant facts about a speaker's spatial location at the time of utterance are what guarantee that the sentence will come out true?\footnote{Although defenders of "metaphysical" versions of analyticity may have views on the matter, I should note that no metaphysical theory of truthmaking or the nature of "facts" need be presupposed here. All we need is a basic grasp of the ordinary practice of calling sentences true in order to see what is at issue. It is correct to call a sentence true only if the world is as the sentence says it is—or, what amounts to the same thing, only if the sentence accurately reports on the facts. Such platitudes are compatible with various deflationary or minimalist accounts of truth, facts, and related notions.}

In order to further motivate the idea that the meaning of certain sentences guarantees (though does not independently determine) their truth, Russell invokes a mathematical analogy. She asks us to consider the basic multiplication function on the natural numbers: ‘\(x \times y = z\)’. In this function, the value of \(z\) is jointly determined by the values of \(x\) and \(y\)—vary either of these values and the value of \(z\) will be affected. Now consider a special subset of the instances of the multiplication function, viz., those in which the \(x\)-argument is assigned the number 0. Russell points out that in these special cases, (which have the form ‘0 \(x y = z\)’), there is a sense in which the value of the \(y\)-argument simply no longer matters—the value of \(z\) will always come out as 0 no matter what value is plugged in for \(y\). The sense in which it doesn’t matter what value we assign to the \(y\)-
argument in this case is supposed to be analogous to the sense in which analytic sentences are guaranteed to come out true no matter what the state of the world is. Following Russell, let’s suppose that the truth-value of a sentence is determined by a function from its meaning together with a state of the world which can be formulated as a set of triples: \(<m, w, v>\). Her idea, then, is that in the case of an analytic sentence, the value of \(w\) (like the value of the \(y\)-argument in multiplication functions of the form \(0 \times y = z\)), can be varied in any number of ways without affecting the truth-value of the sentence because the value of \(m\) is, in Russell’s words, “sufficient, in a sense on its own, to determine the result.”

The key question here is whether this sense of determination is robust enough to avoid the triviality horn of the above dilemma.

Two other recent defenders of a “metaphysical” version of analyticity, Frank Hoffman and Joachim Horvath, share Russell’s general motivations and offer similar sorts of arguments. In order to avoid the dilemma described above, Hoffman and Horvath urge a distinction between two kinds of fact-independence: existence independence and variation independence. They claim that because critics of metaphysical analyticity are guilty of confusing these two distinct types of fact-independence, their chief objections miss the mark. Hoffman and Horvath outline their distinction as follows:

A truth value of a truthbearer is existence-independent of certain facts if it does not depend on the existence of these facts—it is whatever it is, no matter whether these facts exist or not. It is variation-independent of certain facts if it does not change with any possible variation of these facts. These two kinds of independence are not conceptually equivalent, since a truth value can be variation-independent of certain facts without being existence-independent of them.

---

93 Ibid., p. 33.
Armed with this distinction, Hoffman and Horvath characterize analytic truths as those which are true variation-independently—but not existence-independently—of extralinguistic facts. However, they are also keen to point out that

Analytic truth is no exception as far as truthmaking in the existential sense is concerned. It requires the existence of suitable extralinguistic facts, too. But because analytic truth is variation-independent, it does not require the existence of any specific extralinguistic facts. Just about any possible variation of extralinguistic truthmakers will do. So in this properly restricted sense, the ‘factual component’ really does reduce to null, as Quine has put it.95

This maneuver is essentially the same as Russell’s, and helps to further illuminate the sense of determination at work in her proposal. Like Russell, Hoffman and Horvath seek to avoid the Quinean dilemma by articulating a non-trivial, non-absurd, but “properly restricted” sense in which meaning determines the truth of analytic sentences. Analytic sentences are supposed to be special because, being variation-independent of extralinguistic facts, their meanings guarantee that they will always express some truth or other, regardless of how things happen to be in the world.

Perhaps the notion of variation-independence can help demarcate a semantically interesting class of sentences that retain the same truth-value across different contexts, but the sense of truth-value determination at work in the above arguments hardly captures the traditional idea that analytic sentences are true in virtue of meaning alone. If, as these authors claim, truth-making works the same way for analytic sentences as it does for non-analytic ones, we have little reason to think that the analytic-synthetic distinction tracks a real distinction between different kinds of truths or different ways of being true. It might well be the case that in order to be capable of expressing such a broad range of true propositions, a sentence must contain expressions whose reference conditions are satisfied by all possible

states of the world. But it does not follow that such sentences are appropriately characterized as true in virtue of meaning, since all parties agree that meaning does not suffice to explain their truth. Russell, et al might have some reason to think that they have isolated a semantically interesting class of sentences, but they have provided us no reason to conclude that, in spite of Quine’s objections, they have successfully demarcated a metaphysically distinctive class of truths.

On Russell’s view, to say that a sentence is “true in virtue of meaning” is not to say that its meaning alone determines its truth but rather that its meaning determines that it will always express some truth or other. This move does seem to avoid the charge of absurdity, but it is hard to see how it can avoid the triviality charge. For if all a putatively analytic sentence’s meaning determines is that it will express some truth or other—quite possibly a contingent one—then it becomes difficult to see what the philosophical interest of this notion of analyticity is supposed to be. Indeed, it now appears that Russell really has changed the subject; for the sentences to which she affixes the label ‘analytic’ may or may not be a priori knowable, may or may not be meaning-constitutive, and may or may not express necessary truths. Russell’s account is clever and technically impressive, but philosophers concerned with the nature of metaphysical necessity, a priori knowledge, or meaning-constitutivity are unlikely to find much of interest in her account.

2.4 Juhl and Loomis on “Analyticity*”

In their recent book, Cory Juhl and Eric Loomis reconstruct a notion they say is “akin” to analyticity, which they call analyticity*. Juhl and Loomis claim that their notion can do work in the epistemology of mathematics. In particular, they claim that their “proposed notion of analyticity can be used to provide an alternative to the broadly empiricist
justification of basic mathematical principles defended by Quineans.”

However, the first motivating example they give is not a mathematical one, which shows that they take the notion of analyticity* to apply to non-mathematical cases as well. In particular, they ask us to consider the following “best case” example:

Suppose that we say, in the presence of a number of other English speakers, ‘Frenchelors are French bachelors.’ Suppose in addition that those present understand the statement as stipulative. Suppose further that no one has used the term ‘frenchelor’ before this occasion, and everyone agrees to use it to apply to all and only French bachelors. Finally, suppose that everyone takes it for granted that no empirical evidence counts against the statement ‘Frenchelors are French bachelors.’

For Juhl and Loomis, ‘Frenchelors are French Bachelors’ is a paradigm case of analyticity*. It is a sentence that is used stipulatively in order to coin a new term—‘frenchelor’—and Juhl and Loomis make sure to specify that the stipulative act takes place in the presence of “a number” of English speakers. They also note that “everyone” (presumably, all members of the group of English speakers mentioned at the beginning of the passage) agrees to use the word in accordance with the stipulation. Furthermore, all members of the relevant group of English speakers also agree to treat the stipulated statement as empirically indefeasible—that is, they take it that no possible empirical evidence could count for or against its truth.

According to Juhl and Loomis, “When we introduce a stipulation of our particular indefeasible sort into our language, we introduce a coordinative rule concerning some stipulation sentence s” and they claim that this “coordinative rule” has the following form:

(Stip) Sentence s expresses some true proposition p (in our language L). Furthermore, the proposition q, that s expresses a true proposition (in L), is empirically indefeasible. No empirical evidence counts in favor or against the truth of q.

97 Ibid., p. 214.
98 Ibid., p. 218.
Thus, on their view, “When speakers of L accept Stip as a coordinative rule for speaking their language, we say that s is analytic* in L, for speakers of L.”\(^99\) They note that analyticity* requires the empirical indefeasibility of proposition q, but does not necessarily require the empirical indefeasibility of proposition p. On the basis of this observation, Juhl and Loomis propose that there is a special sub-class of analytic* statements for which the stronger requirement holds, and it is in this special sub-class where mathematical statements find their proper home.\(^100\) Thus, they introduce a notion they call “transcendental stipulation”:

\[(T\text{Stip}) \text{ Sentence } s \text{ expresses some true proposition } p. \text{ Furthermore, the proposition } q, \text{ that } s \text{ expresses a true proposition (in L), is empirically indefeasible. Finally, proposition } p \text{ is empirically indefeasible (no empirical evidence counts for or against the truth of } p).\]\(^101\)

Given (TStip), we get a characterization of “t-analytic” statements as those analytic* statements for which (TStip) holds.

Since most stipulation is of the transcendental variety, it is quite rare, Juhl and Loomis claim, to encounter analytic* statements which fall short of being t-analytic, though such cases do exist. As examples, they cite Wittgenstein’s discussion (exploited by Kripke) of the standard meter bar in Paris and Garreth Evans’ ‘Julius’ case.\(^102\) These examples are taken by followers of Kripke to be instances of \textit{a priori} knowledge of contingent matters, and allegedly demonstrate that \textit{a priority} comes apart from necessity. In the Evans example, the sentence ‘Julius is the inventor of the zipper, if there is a unique zipper-inventor’ is supposed to be \textit{a priori} knowable, even though it reports on a contingent fact. The idea here is that although the person to whom we are referring with the name ‘Julius’ actually did invent the

\(^99\) Ibid., p. 219.
\(^100\) Juhl and Loomis treat analyticity* as a property of \textit{statements} rather than a property of propositions or uninterpreted strings with syntactic structure: “A statement (sentence-as-used/understood-on-some-occasion), then, is a candidate for being ‘analytic’ in our present sense.” (p. 218).
\(^101\) Ibid., p 219.
\(^102\) See Kripke (1980), pp. 54-56; and Evans (1982), p. 31.
zipper, it is perfectly possible for that very person to have, for example, died in infancy and thus never invented anything. The point is that Julius might not have invented the zipper, yet it seems that we can know a priori that Julius is the inventor of the zipper, given the way in which we stipulatively introduced the name ‘Julius’. Juhl and Loomis handle this case as follows:

To the extent that J [i.e., ‘Julius is the inventor of the zipper, if there is a unique zipper-inventor’] is understood as an indefeasible stipulation, sentence J is understood to express some true proposition or other. That requirement is what ‘does the work’ with respect to ‘reference fixing.’ However, whatever proposition is actually expressed, on a Kripkean account, is an empirically defeasible proposition. A proposition that states of some individual that he or she invented something is the sort of proposition that empirical evidence counts for or against. Such cases are rare, and this is among the factors that may make them appear surprising. Juhl and Loomis thus suggest that J is analytic*, but not t-analytic. In other words, the proposition that J itself expresses is not empirically indefeasible, even though the proposition that J expresses a true proposition is empirically indefeasible.

Juhl and Loomis go on to claim that unlike traditional analyticity, their notion of analyticity* is to be understood as a kind of response-dependent concept. In particular, they say it is “intention-dependent,” in the sense that “what it is for a statement to be analytic* is to have the linguistic community take it as true and take it as indefeasible.” As a response-dependence account, Juhl and Loomis’s proposal is severely underdeveloped. For one thing, they need to say much more about who counts as the “linguistic community,” given that it is the responses of that community that are doing the work in their account. Why, for instance, think that ‘Frenchelors are French bachelors’ is truly analytic*, given that it is recognized only by a small group of English speakers and not by the rest of “the linguistic community”? Moreover, response-dependent accounts of concepts are commonly taken to specify the

103 Ibid., p. 219.
104 Ibid., p. 226.
105 Ibid., p. 229.
truth-conditions of sincere assertions featuring those concepts. For example, one might hold a view according to which something is funny if and only if the right people are disposed to laugh at it under the right circumstances. One consequence of accepting such an account of the truth conditions of funniness-attributions is that it rules out the legitimacy of Moorean questions like “Sure, the relevant people are disposed to laugh at this joke, but is it really funny?” Analogously, a response-dependent account of what it is to be analytic* would seem to rule out questions of the form “Sure, the community takes statement S to be true and empirically indefeasible, but is it really analytic*?” There would be no room left for dispute.

Then again, perhaps Juhl and Loomis do not intend their response-dependent account to specify the truth-conditions of analyticity* attributions. For they claim that analyticity* is not to be understood as an explanatory notion—i.e., the meaning of an analytic* statement explains neither its truth nor the fact that it is taken to be true—and thus it appears that they would deny that facts about how the community treats certain statements are the truth-makers for analyticity* attributions. Nevertheless, they also claim that analyticity* “can illuminate epistemologically puzzling phenomena such as our apparent non-empirical justification for believing some statements, among other things.” Just what kind of epistemological “illumination” their account is supposed to provide is not clear, however. By “illuminate” they apparently do not mean that a statement’s being analytic* explains how it is knowable a priori along the lines that Boghossian promises. Rather, to say that a statement is analytic* is to say that the linguistic community treats it in a certain way. But this is odd, for whether or not statements are treated as having this status would seem to be an empirical matter. If an extensive sociological survey were to be done, asking all members of the

\footnote{Ibid.}
community whether certain statements are evidentially sensitive to empirical observation or not, then the answer to the question of which claims are analytic* could simply be read off of that data. But surely this is a bad result for Juhl and Loomis, since it is incompatible with the idea that analyticity* attributions are not made true by facts about how the community treats certain statements.

Despite denying that analyticity* is an explanatory notion of the sort that traditional analyticity is supposed to be, Juhl and Loomis qualify this denial by saying that there is a “harmless” and “loose” sense in which meaning “explains” the truth of analytic* statements. For example, if a child asks how we know ‘All bachelors are unmarried’ is true, we can “explain” the truth of this statement by saying “That’s just the definition of ‘bachelor’, it means unmarried man.” The word ‘explain’ clearly does get used in this way. I might, for example, ask someone “Would you explain what you mean by the word ‘culture’ in the context of your political theory?” Or I might tell a child “Let me explain to you what the word ‘mass’ means and how an object’s mass is different from its weight.” However, what seems to be going on in such cases of so-called “explanation” is instead something more like the *inculcating of linguistic know-how* or the *issuing of interpretive licenses*.

On the other hand, paradigm cases of the sort of explanation that empirical science is in the business of providing are, very generally, attempts to single out and describe the operations of underlying physical mechanisms which are causally responsible for certain patterns of observable phenomena. Juhl and Loomis agree that attributions of analyticity* are not explanatory in this stricter, causally-based sense. They are even willing to agree that this concession provides some grounds for thinking that analyticity* is not really a species of analyticity at all. They write: “If one insists that the statement that analytic claims are true in
virtue of meanings is itself analytic*, where the relation ‘in virtue of’ is supposed to be explanatory, we are willing to grant that we are not defending any notion of analyticity.”

Juhl and Loomis arguably read too much significance into Quine’s remarks about so-called “legislative” definitions—i.e., acts of explicit stipulation which introduce novel expressions into the language. On the one hand, Quine seems to make a limited concession to the defender of analyticity when he says that “Legislative postulation … affords truth by convention unalloyed.” But what Quine giveth, he later taketh away: “So conceived, conventionality is a passing trait, significant at the moving front of science, but useless in classifying the sentences behind the lines.” Quine’s point is that when a novel scientific concept is introduced by legislative definition, its status as a “conventional” truth simply will not persist over time and future usage. Conventionality, if there be such a thing, is at best an ephemeral status, not an enduring one. Quine’s so-called “concession” to the defender of analyticity is thus not much of a concession at all. He is simply admitting that we tend to politely go along with legislative postulations in the context in which they are introduced, but that such courtesy does not extend to further contexts of use.

Juhl and Loomis think that, at least in the case of a statement like ‘Frenchelors are French bachelors’, the conventional status can in fact endure through time. They also insists that mathematical statements enjoy this feature, citing examples such as the axioms of set theory, the axioms of arithmetic, and basic equations like ‘2+2=4’. But what about other sorts of statements that have traditionally been counted as analytic? They make a crucial

---

107 Ibid., p. 238.
108 See Quine (1936).
109 Ibid., p. 112.
110 Ibid.
concession to Quine on this score when they admit the class of statements which actually
turn out to be analytic* is actually extremely narrow:

Granting that practically all nonmathematical statements used in the empirical sciences are
empirically defeasible is, on the one hand, a significant concession, for it constitutes a retreat
from accepting many examples that the logical empiricists, and Carnap at every stage of his
philosophy, would have included as examples of analytic statements within empirical science.
But on the other hand, this concession also frees us from having to respond to a wealth of
examples cited by Quineans against the logical empiricists. We think that analyticity* is more
narrowly applied in actual practice than the logical empiricists and Carnap claimed. But we
also think that many explicit stipulations and mathematical statements are clear examples of
non-empirical statements, and that they are employed in accord with conventional norms
that are quite distinct from those governing empirical hypotheses, even highly ‘theoretical’
hypotheses.112

Thus, according to Juhl and Loomis, a statement like ‘Force equals mass times
acceleration’—which was treated as having definitional status prior to Einstein’s
revolutionary contributions to physics—is not now, and never was, analytic*. Even though
that statement may have been taken to be true and indefeasible for a considerable period of
time, Juhl and Loomis must say that because it turned out to be undermined by future
empirical evidence, it was never really analytic*. Moreover, it would appear that, since terms
like ‘whale’ or ‘fox’ are used by biologists, statements like ‘All whales are mammals’ or ‘Every
vixen is a female fox’ will not be analytic* on their account.

That Juhl and Loomis choose to narrow the scope of analyticity* so drastically is
somewhat puzzling, given that they claim to be giving a response-dependent account of the
notion. If to be analytic* just is to be a statement which is treated as true and indefeasible by
the community, then shouldn’t Juhl and Loomis say instead that ‘F = ma’ used to be analytic*,
but no longer enjoys that status? How can the community be wrong about what is analytic*
if having that property just is being treated in the relevant ways? It appears the Juhl and
Loomis wish to have their cake and eat it too: they claim to be giving a response-dependent

112 Ibid., p. 234.
account of analyticity*, but they also claim that nonmathematical statements cannot be analytic*, even if the community in fact treats them as true and indefeasible. Clearly, something has gone wrong with their account.

Recall that Juhl and Loomis promise to “illuminate” the distinctive sort of justification we have for non-empirical statements. When it comes time to say more about what this comes to, however, they fail to deliver. Here is what they say:

So are [analytic* statements] in fact true, known, or justified? Our short answer is that it is not clear that our position requires taking a stand on this matter. One can decide, perhaps arbitrarily, to include them in the extensions of truth, knowledge, or justification. Things will proceed most fluidly in ordinary usage if we do take them to be true, justified, and known, but for careful philosophical purposes, it remains to be shown whether there is a clearly best convention to adopt concerning whether to take analytic* statements to ‘really’ be true, known, or justified, or whether there are good non-conventional theoretical justifications for claiming one or the other.113

This passage is bizarre on its face, given the way Juhl and Loomis bill their project from the start. Like Boghossian’s account of “epistemic analyticity,” the Juhl and Loomis proposal falls far short of the traditional aspirations for an analytic theory of the a priori. Moreover, it is hardly clear what the alleged difference is supposed to be between taking analytic* statements to be true as opposed to “really” true. What would it be for a statement to be “really” true, rather than just plain old true?

Juhl and Loomis make much of the fact that working mathematicians treat mathematical axioms differently from empirical statements. It is in light of this fact about mathematical practice that they claim that such axioms have a distinctive epistemic profile, such that empirical evidence is not relevant to their justificatory status. Juhl and Loomis do not seem to appreciate, however, that bestowing this kind of distinctive epistemic profile on certain sentences happens in non-mathematical contexts too, and may be motivated by a

113 Ibid., p. 230.
variety of different background goals. A group of mariners, for example, might hold fast to
the sentence ‘A whale is a fish’, despite being aware that biologists reject it on the grounds
that whales have lungs and give birth to live young. For by the lights of these mariners, being
a fish has everything to do with being an aquatic animal resource of a certain kind and
nothing to do with the academic classifications of biologists. As such, instead of thinking
of the special status as something that is instituted by the responses of “the linguistic
community” (whoever that is supposed to include), it would be more fruitful to focus on
what goes on in particular conversational contexts where a group of interlocutors are
engaged in the task of trying to interpret one another and share information.

What matters in actual conversation is that there be, at any given point in the
discourse, certain interpretive fixed points provisionally agreed upon by the participating
interlocutors which provide pro tanto standards for distinguishing changes of meaning from
changes of belief. For those purposes, the notion of analyticity* is of little help, since the
vast majority of sentences that clearly play this interpretation-guiding role will not count as
analytic*, given how narrow the class of such statements apparently is. At the end of the day,
the list of statements which count as genuinely analytic* includes ‘Frenchelors are male’ and
‘2+2=4’, but fails to include the vast majority of sentences that ordinary competent speakers
would be inclined to sort into the ‘analytic’ pile. Once again, we have been presented with an
ersatz notion—analyticity*—which turns out to be of little philosophical interest.

114 See Donnellan (1962), and Sidelle (2007). See also Putnam (1962b).
2.5 Concluding Remarks

It appears that in order to give an account of analyticity which avoids the most serious objections to the traditional notion, one must change the subject and give an account of something else instead. Upon close examination, each of the three most worked-out recent attempts to resurrect analyticity present us with something significantly different from the traditional notion. In each case, we are assured at the start that the notion of analyticity to be articulated and defended will do at least one of the things that traditional analyticity was supposed to do. Boghossian says his notion can account for the *a priori* status of logic and conceptual truth in the “strong” (i.e., empirically indefeasible) sense. Russell says she can make substantive sense of the idea of “truth in virtue of meaning.” Juhl and Loomis promise that their account will “illuminate” the distinctive kind of warrant we have for non-empirical claims. However, as we have seen, it is doubtful that any of these authors actually succeed in articulating a notion of analyticity that both avoids the major objections and also makes good on even one of analyticity’s three philosophical roles. This may suggest that Quine was right to question the theoretical utility of any notion that deserves the name ‘analyticity’.

Nevertheless, strong intuitions remain that there must be something different or special about many of the sentences traditionally classified as analytic. Intuitively, these sentences are such that to deny them is not merely to misdescribe the world, but to sin against meaning. Yet, even sentences which turn out to be false on empirical grounds can have this status, at least for a time. If we think in terms of analyticity, we’ll be inclined to say that, if such a sentence turns out false, then it was never analytic, and anyone who treated it as analytic was simply mistaken. But if we stop thinking in terms of analyticity, we can stop wondering which sentences are genuinely analytic and instead focus on the following
question: Why might it be useful for language-users to employ predicates like ‘means the same as’, ‘is synonymous with’, ‘is analytic’, ‘is a definition’, etc., even if no sentences are actually analytic in anything like the traditional sense?
Chapter 3
Constitutivity Commitments and the Pragmatics of Conversation

3.1 Introduction

To be a competent speaker of a language, one surely needs to be able to discriminate between truth and error. But competence also requires something more: it is essential that the competent language-user be able to tell the difference between pieces of behavior that do, and pieces behavior that do not, count as normatively significant moves in a public language-game. In other words, a competent speaker needs a sense of what is constitutive of the very activity of speaking her language, much as a competent chess player needs a sense of what is constitutive of the activity of playing chess. In chess, as in conversation, there are good moves, there are bad moves—and then there are non-moves: pieces of behavior that are incompatible with continued participation in the game. One can make lots of bad moves and still count as playing the game, but make too many non-moves and your status as a genuine participant comes into question.

[115 See Midgley (1958) and Searle (1969) for more on this use of ‘constitutive’. As Fennell (2013) puts it, “Constitutive norms are those that one cannot contravene in the main and still be engaged in the practice in question” (p. 63). Constitutive norms are to be distinguished from both categorical and hypothetical imperatives, as the force of constitutive norms does not depend on the having of certain desires, and constitutive norms are not automatically binding on all rational agents. Rather, they are norms that establish what sorts of behaviors or physical events shall count as genuine moves within a certain social practice or activity.]
There are, of course, differences between playing chess and speaking a natural language. The rules of chess provide a standard of correct play which uniquely determines the class of admissible moves relative to any possible state of the board. Things are different with respect to natural language. For although there is considerable linguistic conformity among speakers with similar cultural backgrounds, there is no official list of rules from which the answers to interpretive questions can be read off. Competent speakers, by virtue of their competence, have a partial say in what constitutes correct usage; but insofar as the rules of chess are conceived of as static, competent chess players lack any analogous sort of authority. Each language-user has her own sense of what is constitutive of the activity of speaking her language, but if she is rational and cooperative, she will generally work to make her linguistic dispositions mirror those of her peers.

In this chapter, I give an account of what I call constitutivity commitments—the crucial interpretive dispositions language-users need to acquire in order to count as genuine participants in a linguistic practice. My account provides a pragmatic explanation what we are doing when we engage in disputes over a word’s meaning, and I argue that the same pragmatic mechanism also explains metaphysical disputes over a thing’s essence. I contend that undertaking constitutivity commitments is part of what it is to be a rational, cooperative participant in a discourse interaction, and our ability to assess, negotiate, and coordinate these commitments is what makes mutual understanding, and thus successful communication, possible.
3.2 From Analyticity to Constitutivity

A competent interpreter needs the ability to tell whether her conversational partner does or does not mean what she means by various expressions in various contexts. In cases of significant interpretive mismatch, the interlocutors are in an important sense playing different games—speaking different languages—and thus talk past one another. To avoid such communication breakdowns, the parties need to be in rough agreement about which behaviors are to count as genuine moves (good or bad) within the game, and which behaviors constitute evidence that someone doesn’t grasp the rules, or is playing a different game with different rules. These interpretive standards may or may not persist across different contexts or conversations, but the important point is that the interlocutors must roughly agree on some such standards at any given point in a discourse. Without sufficient alignment of the interlocutors’ interpretive dispositions, it becomes impossible for information exchange to take place.

For an example of how this kind of agreement is achieved, consider the activity of linguistic instruction. If I wanted to teach someone the meaning of the word ‘octagon’ without a visual aid, it would be natural to make use of the sentence ‘An octagon is an eight-sided polygon’. And if I were to hear someone talking about octagons in a way that seemed to imply the negation of that sentence, it would be natural to have doubts about whether the speaker means what I mean by the expression ‘octagon’ or, alternatively, whether the speaker even understands what an octagon is. Of course, language learning does not always or even typically involve this sort of explicit definition. We typically learn about basic geometrical figures by being presented with a range of appropriately shaped objects or

---

116 Depending on the situation, I might also have doubts about what my interlocutor means by ‘polygon’ or ‘eight-sided’.

67
graphical representations, thereby acquiring a disposition to associate tokenings of the word with relevantly similar objects. Indeed, most of our words are learned by observing the way others apply various expressions and extrapolating from these observed patterns of application a sense of how to use them correctly.

But whichever way a person ends up learning a particular word, there will always be certain sentences and inferences featuring that word which she is disposed to treat as especially relevant in contexts where interpretive coordination is the primary task at hand. A sentence or inference that plays this role does so in virtue of being treated as a sort of interpretive guideline—a pro tanto standard by which we assess whether our conversational partners are using certain expressions with the same senses that we use or would use them. In typical cases, sentences which play this interpretive role have a special status: they are simply taken for granted as obvious conceptual truths to which any competent speaker should assent without hesitation.

Sentences like ‘Every vixen is a female fox’ and ‘No bachelor is married’ have historically been classified by philosophers as analytic because they appear to be a priori knowable and seem to be true by virtue of meaning alone, no matter what the worldly facts are like. But even if a viable analytic-synthetic distinction could somehow be drawn, it appears that intuitively meaning-constitutive principles are not always analytic truths. For example, as Hilary Putnam has pointed out, before the development and widespread acceptance of Einstein’s relativistic physics, sentences like ‘Kinetic energy is equal to one half the product of mass and velocity squared’ and ‘If two straight lines are perpendicular to a third then the two do not meet’ were widely agreed to be meaning-specifying definitions of ‘kinetic energy’ and ‘straight line’. Later, as relativistic physics became widely accepted, these
sentences were rejected as false, and thus lost the special status they previously enjoyed. Cases like this can be found throughout the history of science. They are cases in which certain sentences play the role of interpretation-guiding principles for a significant period of time, but are later discovered to be false in light of new evidence and/or revolutionary changes in scientific theory.

Perhaps such considerations suggest that a reconceptualization of the notion of analyticity itself is in order. That is, we might opt for an explication according to which analyticity is to be understood as a sort of temporary privileged status we may choose to grant to certain sentences, but which may later be rescinded. In fact, Friedrich Waismann seems to hold a view of this kind. In the second of a series of six articles on analyticity from the early 1950s, Waismann writes:

[Waismann writes:] 

[It is significant that we do not only “find out” that a given statement is analytic; we more often precisify the use of language, chart the logical force of an expression, by declaring such-and-such a statement to be analytic. … It is precisely because, in the case of ‘analytic’, the boundary is left open somewhat that, in a special instance, we may, or may not, recognize a statement as analytic.]

Waismann suggests that analyticity isn’t so much a property that we discover certain sentences to have, but rather a property that we declare those sentences to have. The scare quotes around the phrase “find out” seem to indicate that, on Waismann’s view, we don’t detect the property of analyticity, we project it. He also notes that the extension of the predicate ‘is analytic’ is importantly open in the sense that what counts as analytic may change over time.

---

117 Putnam (1962a), pp. 42-50. Arguably, however, the expressions ‘kinetic energy’ and ‘straight line’ did not thereby change their meanings. For although earlier theorists had certain beliefs about energy and straight lines that turned out to be false, it does not follow that those theorists were not talking and thinking about energy and straight lines. See also Putnam (1970).


119 I make no strong claims about the historical accuracy of this interpretive suggestion (though there is virtually no secondary literature on Waismann to come into conflict with). It may be noted, however, that Hume-inspired expressivist/non-cognitivist views were quite common among Waismann’s logical empiricist peers. See Kraut (forthcoming-a) for similar suggestions about an expressivist interpretation of Carnap on ontology.
This openness would help explain what is going on in Putnam-style examples like those mentioned above.

Another significant gesture toward some kind of projectivism about analyticity occurs in the following passage from A. J. Ayer:

We have already explained how it is that these analytic propositions are necessary and certain. We saw that the reason why they cannot be confuted in experience is that they do not make any assertion about the empirical world. They simply record our determination to use words in a certain fashion. We cannot deny them without infringing on the conventions which are presupposed by our very denial, and so falling into self-contradiction.\footnote{Ayer (1936), p. 84.}

It is important to distinguish two claims Ayer makes here. First, he claims that when you assert an analytic sentence, you aren’t \textit{saying} anything about the world—that is, analytic indicatives are non-descriptive or non-factual. Second, he claims that what you are \textit{doing} when you make such an assertion is expressing your intention to use words in a certain way—that is, the purpose or function of asserting analytic indicatives is to express our commitment to certain linguistic conventions or rules of use. Note that Ayer’s two claims are logically independent in the sense that neither follows directly from the other.

I think that these passages from Ayer and Waismann contain some important insights about the purpose or function of the expression ‘analytic’ and related vocabulary. By explicitly declaring a sentence to be analytic, a speaker is able to convey pragmatic information about her interpretive commitments, and thus make explicit her sense of how certain bits of language are to be appropriately interpreted. This means that even if Quinean critiques of analyticity successfully show that the notion cannot explain necessary truth or ground \textit{a priori} knowledge, it does not follow that analyticity-attributions lack any legitimate function in discourse.
Linguistic interpretation is a highly complex social-coordinational task. Communication breakdown is not uncommon, and interpreters often need to make decisions about whether an interlocutor’s utterances involve legitimate extensions of usage, unacceptably deviant usage, or simply amount to a change of subject. How we are disposed to make these interpretive decisions depends upon the interpretive commitments that we have undertaken; and since successful communication requires that interlocutors coordinate their interpretive decisions, we need to be able to keep track of one another’s interpretive commitments. In order to keep track of these commitments, make public note of them, and present them for critical scrutiny, we need some vocabulary with which to express them. As such, insofar as it allows us to make our interpretive commitments explicit, the philosophers’ technical term ‘analytic’ seems quite useful to have around. Much the same could be said for ordinary, non-technical talk of synonymy, definition, and word-meaning.

Unfortunately, the technical term ‘analytic’ remains problematic in ways that ordinary talk of definitions or synonymy is not. In particular, the notion of analyticity is problematic because it is factive. i.e., the claim that sentence \( S \) is analytic entails the claim that sentence \( S \) is true. The trouble with this is that a sentence need not be true in order to play the role of an interpretation-guiding constitutive principle. Sentences that play this interpretive role are such that, relative to a particular conversational context, it is prima facie inappropriate to deny them. But this is entirely compatible with such sentences turning out to be truth-valueless or even false—in which case, they certainly won’t be analytic truths.

Because interpretation-guiding principles need not be true, we ought therefore to stop courting confusion and simply drop the terminology of ‘analyticity’ altogether. Since the goal is to capture what is distinctive about sentences that play the interpretive role, and since
sentences can and do play this role even when they aren’t true, there is good reason to seek out alternative vocabulary with which to characterize them. Moreover, what is most distinctive about sentences that play the interpretive role is the way in which they reflect a competent speaker’s sense of what is constitutive of the activity of speaking her language. The terminological choice is thus clear: I shall henceforth eschew the term ‘analytic’ and exclusively refer to these interpretation-guiding sentences as constitutive principles.

Since constitutivity is not a factive notion, it is perfectly possible for an expression to be governed by a set of constitutive principles which are inconsistent. Constitutive principles may also be schematic in form, and so neither true nor false until contextually fleshed out. Given that analytic sentences are supposed to be true, constitutivity should not be thought of as a species of analyticity. Instead, we should think of constitutivity as a kind of successor notion, suitable for theorizing about interpretation and linguistic understanding, but orthogonal to the traditional explanatory projects most closely associated with the analytic-synthetic distinction.

It should be noted that treating sentence $S$ as a constitutive principle is not a matter of simply believing $S$ in the same way that one believes ordinary non-constitutive sentences. A person who merely believes ‘Bachelors are unmarried males’—but does not think that this sentence has any kind of special status which would distinguish it from the non-constitutive sentence ‘Bachelors tend to enjoy action movies’—seems intuitively to betray a lack of

---

121 It is for this reason that I choose not to follow Jamie Tappenden in calling such sentences “pre-analytic”. See Tappenden (1993).
122 Some of those who advocate “meaning-inconsistency” approaches to semantic paradoxes—for example Eklund (2002), (2007) and Scharp (2005), (2007), (2013)—employ this terminology in a very similar spirit. See also Davidson (1970), where it is claimed that “the whole set of axioms, laws, or postulates for the measurement of length is partly constitutive of the idea of a system of macroscopic, rigid, physical objects” (p. 221).
123 Constitutive principles governing gradable adjectives like ‘tall’ or ‘big’ are arguably like this. See Barker (2002).
competence (or at best an incomplete competence) with the word ‘bachelor’. Another way to put what amounts to the same point would be to say that the former sentence is especially competence-relevant because it seems to specify the essence of bachelorhood, whereas the latter sentence simply reports on an accidental property of some, but not all, bachelors.

A further complication arises from the fact that it is not always necessary for a person to actually believe the relevant constitutive principles in order to count as linguistically competent with a certain expression. For example, non-racists can surely understand racially pejorative expressions, even though they do not believe that the constitutive principles governing such expressions are true. Consider the classic example of the pejorative expression ‘boche’. The constitutive principles governing ‘boche’ are (1) if \( x \) is German, then \( x \) is boche; and (2) if \( x \) is boche then \( x \) is cruel. Together, these principles imply an empirically false generalization: namely, that all Germans are cruel. Note, however, that one can avoid becoming committed to this false generalization by simply avoiding the use of the expression ‘boche’. But in order to deliberately avoid using ‘boche’, one must be capable of understanding it when the expression is used by someone else. After all, how else would one know what to avoid?

The sincere and habitual user of the term ‘boche’ takes himself to be entitled to conclude of any arbitrary German that she is cruel; but from the point of view of the non-racist, that inferential entitlement is defeated by the empirical facts. Nevertheless, the non-racist understands perfectly well what constitutes using ‘boche’ in accordance with its meaning, even if she has no disposition to employ (i.e., use rather than merely mention) the

---

term herself. Indeed, it is precisely because she understands the expression that she is in a position to stand in moral opposition to its use by others in her speech community.

In light of such considerations, we ought to say that to understand expression $E$ is to know what one’s dispositions would have to be like in order to use $E$ correctly. But knowing how to use $E$ correctly is not simply a matter of being reliably disposed to apply it to whatever it is true of and withhold its application otherwise. Rather, knowing how to use $E$ correctly is a matter of knowing that (a) there are certain inferential transitions that count as free moves prima facie available to any competent $E$-user, and (b) there are certain pieces of $E$-behavior that tend to indicate that someone is playing a different game (or no game at all). We should thus understand an individual’s sense of what counts as correct $E$-usage in terms of the sentences and inferences she treats as having the status of constitutive principles with respect to $E$.

### 3.3 Constitutivity and Verbal Disputes

For each question on the following list, imagine a dispute in which the parties agree on all the relevant background facts about ordinary observable matters, yet one disputant answers “yes” while the other answers “no”:

1. Is a drinking glass a cup?
2. Is any cocktail served in a V-shaped glass a martini?
3. If a man and a squirrel simultaneously circle a tree, both facing the trunk, does the man go around the squirrel?

\[125\] The account I develop here is committed to what Fennell (2013) calls the contextualist reading of the thesis that meaning is normative: “In all contexts of a term’s usage, there is some use or belief (or set of uses/beliefs) or other, possibly different, that is (are) constitutive of its meaning what it does in that context” (p. 66).
(4) Is Secretariat, the record-setting racehorse, an athlete?

(5) Is a whale a fish?

(6) Is Pluto a planet?

(7) Is the Affordable Care Act a socialist policy?

(8) Is Egypt a democracy?

(9) Is a civil union between same-sex partners a marriage?

(10) Is a second-trimester human fetus with a rudimentary capacity to feel pain a person?

(11) Is a semi-automatic rifle with a pistol-grip and a twenty-round magazine an assault weapon?

(12) Is an individual with an annual income above $250,000 rich?

Intuitively, disputes over questions like these seem to be verbal, rather than substantive. That is, disputants who give incompatible answers do not seem to be disagreeing about empirically decidable matters, but rather about the proper use of certain linguistic expressions (e.g., ‘cup’, ‘martini’, ‘go around’, ‘athlete’, ‘fish’, ‘planet’, ‘socialist’, ‘democracy’, ‘marriage’, ‘person’, ‘assault weapon’, and ‘rich’). In such cases, it seems that if only the disputants could agree on how to define their terms, the dispute would dissolve and the background agreement between the parties would be revealed.

In practice, many such disputes often take the form of disputes over essences or natures—the nature of personhood, for example, or the essence of democracy. To frame a

---

126 Question (1) comes from Hirsch (2005), (2) comes from Bennett (2009), (3) comes from William James as cited by Chalmers (2011). Question (4) comes from Ludlow (2006). Questions (5)-(8) are taken from and/or inspired by Sidelle (2007). Question (9) was inspired by issues discussed in Stivers and Valls (2007). Question (10) was inspired by issues discussed in Stivers and Valls (2007). Question (11) was inspired by issues discussed in Stivers and Valls (2007). Question (12) was inspired by issues discussed in Stivers and Valls (2007). Question (12) comes from reflection on issues raised by Warren (1973) and Sumner (1987). Question (11) was derived from considerations raised in a 2012 Slate.com article, “The Gun Glossary”, by Mark Joseph Stern: (http://www.slate.com/articles/news_and_politics/explainer/2012/12/the_gun_glossary definitions_of_firearm_lingo_and_types_of_weapons.html). Question (12) was inspired by recent American political debates over tax policy.
dispute in terms of essence is to frame it as a dispute over, as John Locke puts it, “the very being of [the] thing, whereby it is what it is.” 127 It is not at all clear, however, that the use of such vocabulary sheds any explanatory light on the situation, because essence-talk itself cries out for explanation. What is wanted, in particular, is an explanation of why creatures like us develop such vocabulary in the first place and why its employment persists. To this end, there are broadly speaking two sorts of explanatory strategy available. 128 On the one hand, one might hold that the discourse in question (essence-talk) exists primarily in order to allow us to accurately represent genuine features of the world, and that its predicates express real properties—where we understand this to mean, roughly, that the discourse is about, and assertions within it are made true by, explanatorily ineliminable essence-properties. On the other hand, one might hold that essence-talk is quite useful in certain circumstances, but when explanation is the order of the day, talk of essence is ultimately eliminable. There are, of course, many subtly different ways to implement either sort of strategy, but this rough characterization captures the explanatory contrast I mean to emphasize. Proponents of the former type of strategy are unified by a commitment to the explanatory ineliminability of essence, while proponents of the latter strategy are unified by a rejection of that commitment. I will pursue the latter sort of strategy. 129

The fact that essence-talk is so often deployed in such disputes is quite understandable, given the close ties between essence and meaning. As Quine famously wrote, “Meaning is what essence becomes when it is divorced from the object of reference

127 Locke (1690), Book III, Chapter III, §15.
128 I deliberately avoid the labels “Realism” and “Anti-Realism” here because of the distracting and potentially misleading philosophical baggage associated with them.
129 See Kraut (2010) for a discussion of this sort of strategy as it applies to property-talk.
and wedded to the word.”

More recently, Kit Fine has pointed out “systematic analogies” between necessity and analyticity, and also between essence and meaning. Fine writes the following:

We have seen that there exists a certain analogy between defining a term and giving the essence of an object; for the one results in a sentence which is true in virtue of the meaning of the term, while the other results in a proposition which is true in virtue of the identity of the object. However, I am inclined to think that the two cases are not merely parallel but are, at bottom, the same. … A definition, on this view, would therefore state an essential property of [the meaning of] the word.

As Fine might put it, necessary truths seem to be grounded in facts about essence, while analytic truths seem to be grounded in facts about meaning. Hence, just as necessity flows from essence, analyticity flows from meaning. Fine notes that definitions are statements of meaning which specify the essential properties an object must possess in order to properly fall under the extension of the term being defined. This further demonstrates the extremely close relationship between knowing the meaning of a word and grasping the essence of the thing denoted by that word. Such considerations lead Fine to conclude that “the activities of specifying the meaning of a word and of saying what an object is are essentially the same.”

Whether they are conducted using the vocabulary of essence or not, there are many verbal disputes that appear to be pointless and a waste of time. Typical disputes over questions (1)-(4) seem to have this sort of trivial air. But many other verbal disputes—even when it is clear to everyone that they are verbal—are extremely important because of the potential practical, social, or political fallout deciding them one way or the other might have. Questions (5) and (6) have significant practical consequences for the classificatory practices

---

130 Quine (1951), p. 22.
133 According to Ludlow (2006), the example in (4) comes from a real debate he heard on sports talk radio. Presumably some of the disputants—those who take sport-related issues very seriously—might balk at my suggestion that a dispute over (4) has a trivial air.
of zoology and astronomy, and questions (7)-(12) have obvious moral and/or socio-political import. In some cases, disputes are only partly verbal, but the verbal element of the dispute masks or obscures an underlying substantive disagreement, making it impossible for the parties to identify—let alone settle—their genuine differences. For all these reasons, the ability to diagnose and resolve verbal disputes is a crucially important skill for any rational, language-using creature.

Rudolf Carnap famously argues that many of the disputes that exercise philosophers—especially disputes in metaphysics—are actually verbal disputes.\(^\text{134}\) Although Carnap thought he could use analyticity to explain necessary truth and \textit{a priori} knowledge, his deepest motivations for embracing the analytic-synthetic distinction are arguably based on considerations associated with interpretation.\(^\text{135}\) According to Carnap, intractable metaphysical disputes arise because the disputants talk past one another. More precisely, the conflict is unresolvable because the disputants fail to realize that they each endorse different sets of analytic meaning postulates and thus each attach different meanings to philosophically significant expressions.

My account is intended to preserve the spirit of Carnap’s view, but do so in a way that is compatible with Quinean skepticism about analyticity. Like Quine, I do not think that analyticity claims are explanatory. Being told that a certain sentence is analytic does not constitute an explanation of that sentence’s modal or epistemic attributes. There is, however, an important pragmatic function that underlies the making of analyticity-claims (as well as synonymy-claims and explicit definitions). Moreover, the same pragmatic function underlies

\(^{134}\) See Carnap (1950), where he argues that the debate between platonists and nominalists over the existence of numbers is a verbal dispute because the respective parties mean different things by expressions like ‘number’ and ‘real entity’ and thus hopelessly talk past one another.

\(^{135}\) Carnap’s emphasis on the importance of interpretive coordination underpins his career-long crusade against “pseudo-problems” in philosophy. See Carnap (1963), pp. 43-44.
paradigmatically metaphysical discourse regarding essential properties, identity-conditions, persistence-conditions, and existence-conditions. The common pragmatic function served by both meaning-talk and essence-talk is that of expressing constitutivity commitments. More precisely, such claims serve to (a) publicly indicate the interpretive dispositions of the speaker and (b) facilitate the normative assessment, and thus communal alignment, of such dispositions. Typical metaphysical disputes are thus verbal disputes on this view, in the sense that they are to be explained pragmatically as clashes of incompatible interpretive commitments.

Before elaborating further on my own view, however, I want to consider an illuminating dialectic that has emerged in response to Eli Hirsch’s broadly neo-Carnapian account of metaphysical disputes about ordinary physical objects. Hirsch’s focus is on debates between mereological essentialists (i.e., those who share Roderick Chisholm’s intuitions about the identity, existence, and persistence conditions of physical objects) and four-dimensionalists (i.e., those who share David Lewis’s intuitions about the identity, existence, and persistence conditions of physical objects). Hirsch argues that such disputes are verbal rather than substantive. On Hirsch’s account, verbal disputes occur when the parties agree on all the underlying facts about the world, but disagree about how to interpret certain ordinary language expressions. The participants in the debates in which Hirsch is interested seem to disagree about certain metaphysical claims—e.g., claims about material object composition, colocation, identity, persistence, etc. However, according to Hirsch, the parties don’t really disagree about substantive facts, but instead merely disagree about what

136 See Hirsch (2005), pp. 75-82 for details.
the quantifier expressions mean. As a result of this meaning-divergence, different
metaphysicians assign different truth conditions to the ontological sentences in question.\textsuperscript{137}

So described, disputes about the metaphysics of ordinary physical objects exemplify
what Hirsch calls the “simplest paradigm” of a verbal dispute:

The simplest paradigm of a verbal dispute—the simplest way it can happen that each side of
a dispute can find a charitable interpretation that makes the other side come out right—is
where, for each disputed sentence \( D \), there are two undisputed sentences \( U_1 \) and \( U_2 \), one
true and one false, such that one side holds that \( D \) is (a priori necessarily) equivalent to \( U_1 \)
and the other side holds that \( D \) is equivalent to \( U_2 \). Each side can then assign charitable
truth conditions to \( D \) in the other side’s language simply by assuming that in that language
the other side’s asserted equivalence holds.\textsuperscript{138}

For Hirsch, then, a verbal dispute over \( D \) is a disagreement about which sentence—\( U_1 \) or
\( U_2 \)—\( D \) is “a priori necessarily” equivalent to. In other words, it is a dispute about whether \( D \)
is \textit{analytically} equivalent to \( U_1 \) or \( U_2 \).

Karen Bennett has raised an important objection to Hirsch’s account of verbal
disputes as it applies to debates in metaphysics. Consider the ontological question of
whether composite physical objects exist. Call the metaphysician who answers “yes” the
\textit{Believer} and the metaphysician who answers “no” the \textit{Nihilist}. According to Bennett:

The linking principle ‘if there are simples arranged F-wise in region R, there is an F in R’ is
not analytic in the language of the believer in composite objects. The key piece of my
argument for this claim is the simple fact that the believer does not think that the
composites are identical to anything that the nihilist accepts. When she says that she believes
in tables, she is saying that she believes in tables that are \textit{numerically distinct} from the
simples.\textsuperscript{139}

Bennett’s objection can be summed up as follows. In order for a dispute between two
metaphysicians to be verbal, it must be the case that one party thinks that the disputed

\textsuperscript{137} Hirsch thinks that if a dispute is verbal, the only thing for the disputants to be right or wrong about is which
side’s language “is (closest to) plain English” (p. 70). The assumption is that ordinary competent speakers (in,
e.g., Britain, America, or Australia) speak a common determinate language called “plain English”. See Ludlow
(2006) for a litany of compelling reasons against thinking of natural language in this way.

\textsuperscript{138} Ibid., p. 83.

\textsuperscript{139} Bennett (2009), p. 54.
sentence $D$ is analytically equivalent $U1$ while the other party thinks that $D$ is analytically equivalent to $U2$. But metaphysicists do not typically take these equivalences to be analytic. Since the equivalences are not (taken to be) analytic, $U1$ and $U2$ fail to specify what $D$ means in the respective languages of the two sides. Therefore, it is not the case that typical disputes in metaphysics are verbal disputes.

In defense of a broadly Neo-Carnapian approach to such disputes, Alan Sidelle responds to Bennett’s argument by distinguishing “interesting” verbal disputes from trivial or uninteresting ones. According to Sidelle, interesting verbal disputes occur “when both of the candidate meanings are plausible candidates for the meaning of the public term, and it doesn’t seem obvious that the public term is simply ambiguous.”

Sidelle thinks that the ontological disputes Bennett takes to be substantive are actually verbal disputes of the interesting variety. He writes:

In [interesting] cases, I suggested, the public word(s) in question will be such that the various candidates are all viable interpretations. When this is so, it will not be plausible to definitely assign one of the meanings even to one of the speakers. And so, even though the relevant claims will be analytic on the given interpretations, they won’t be analytic—not determinately. But this doesn’t make the dispute any more genuine. I think that in cases of this sort, it is extremely natural for the parties involved to view the different interpretations as different substantive theories about the nature of $F$—for the very reason that (a) they seem to be determinately fixed on a common subject matter, but (b) because of the dispute, neither seems to be analytically true of $F$. But these assumptions are founded, I think, on a naïve sensibility of one’s grasp of one’s language, and failing to appreciate the challenges for determinacy.

Sidelle’s point here is that interesting verbal disputes typically appear substantive to the disputants because the disputed sentences feature expressions whose meaning is indeterminate. To say that the meaning of a linguistic expression is indeterminate is to say that the conditions for its correct application (and the inferential consequences of its

---

141 Ibid., pp. 107-108.
application) are unfixed or unsettled. If the meaning of an expression \( E \) in sentence \( D \) is indeterminate, then there is no fact of the matter about which of the two interpretations, \( U_1 \) or \( U_2 \), is the uniquely correct interpretation of \( D \). Thus, according to Sidelle, the indeterminacy of \( E \) is what explains why neither \( U_1 \) nor \( U_2 \) appears analytically equivalent to \( D \). In the kinds of cases Bennett highlights, the disputants take themselves to be offering incompatible substantive theories about the essence or nature of a common subject matter, but in light of meaning-indeterminacy, their dispute is best explained as a dispute about correct interpretation—i.e., a kind of verbal dispute.

Recall that the Believer in composite objects thinks that it is appropriate to say ‘There is a table in region \( R \)’ when and only when it would also be appropriate to say ‘There are simples arranged table-wise in region \( R \).’ The Nihilist, on the other hand, insists that it is never appropriate to say ‘There is a table in region \( R \)’, even under circumstances in which it is appropriate to say ‘There are simples arranged table-wise in region \( R \).’ Although the disputants are not willing to call them analytic, the relevant linking principles are taken to be metaphysically necessary and are not claimed to be justified on empirical grounds. The Believer might well say that it lies in the nature of tables that a table is numerically distinct from the simples that compose it, but the Believer will deny that it is analytic that the table is distinct from the simples.

Once again, the notion of analyticity is muddying the waters here. The fight between the Believer and the Nihilist is not best understood as a disagreement over which linking principles are genuinely analytic. Rather, it is best explained as a clash of conflicting

---

142 For a wide variety of compelling reasons to take the phenomenon of meaning indeterminacy seriously, see Goodman (1955), Quine (1960), Kripke (1982), Wilson (1982), (2006), and Price (2009). Taking indeterminacy seriously is compatible with holding that there are clear-cut cases of correct and incorrect linguistic usage, so long as one admits that nothing about our past usage, intentions, dispositions, etc., uniquely determines what counts as correct usage in all possible situations.
interpretive stances—i.e., a disagreement over which principles ought to be treated as constitutive. But although it can be characterized as a kind of verbal dispute, there is no reason to think that it is therefore not a genuine dispute. Even if, as Hirsch suggests, the Believer and the Nihilist can charitably interpret each other so that the ontological claims of each side come out true in their respective languages, I contend that there is still a genuine pragmatic conflict between them over how to use words properly. This is what underlies typical disputes over questions like (1)-(12), and, I contend, it is what is at stake in typical metaphysical disputes.

It should be noted that what distinguishes the “uninteresting” or trivial verbal disputes from the more “interesting” ones is not just that the alternative interpretations are, as Sidelle puts it, “plausible candidates” for the meaning of the expression in question. Rather, the way cases get sorted—whether they are deemed intuitively interesting or intuitively trivial—depends on the background interests and goals of the participants in the relevant disputes. For example, it simply doesn’t matter very much to most ordinary people whether a drinking glass counts as a cup—in light of a whole slew of (contingent) facts about the cultural role of drinking vessels, there are no weighty practical consequences associated with this sort of dispute. On the other hand, it matters tremendously to many people whether, for example, a borderline-sentient fetus counts as a person. The dispute matters, in this case, because it addresses a social policy issue with significant practical consequences about which ordinary people tend to have strong moral convictions. Likewise, it clearly matters to metaphysicians (though not to ordinary speakers) whether a collection of

143 I am not aware of any actual human culture in which great practical significance attached to the question of what counts as a cup. However, under the right cultural circumstances—where, say, the material from which cups are made has profound religious significance—there very well could be lots at stake in such a dispute.
simples-arranged-table-wise counts as a table, and this helps explain why they take themselves to be having a substantive, non-verbal dispute over the nature and existence of tables.

In general, it may be observed that when disputants truly care about an issue and take themselves to have a real stake in the matter, they tend to view the dispute as a deep one about the nature or essence of the subject matter in question. On the other hand, when the disputants see little or no significant practical upshot associated with one answer or the other, they tend to think of the dispute as a trivial one about “mere” words. In those cases where the disputants have a significant practical stake in the matter, they tend to indicate this by conducting the dispute at the object language level (employing, e.g., talk about nature or essence). In those cases where the disputants have little or no practical stake in the matter, they tend to indicate this by conducting the dispute at the metalinguistic level (employing, e.g., talk about meanings or definitions).

An important class of exceptions to this general observation should be noted, however. There may be contexts in which the disputants explicitly and consciously take themselves to be negotiating or litigating the meaning of a certain expression, and yet also take the dispute to have important practical consequences. We might imagine, for example, judges or lawmakers with no inclination to employ essentialist locutions who engage in exclusively metalinguistic disputes over how to define terms like ‘assault weapon’, or ‘rich’, yet still take those disputes to have extremely important practical consequences. It is usually, however, simpler and more convenient to conduct such disputes at the object-language level by making de re modal claims (‘Assault weapons are necessarily fully-automatic’) or essentialist claims (‘Being fully-automatic is essential to being an assault weapon’).
Of course, claims about what is essential to being a cat do not have the same truth-conditional content as claims about the meaning of the expression ‘cat’. The claim ‘Cats are essentially felines’ is about cats, while the claim ‘The word ‘cat’ means a type of feline’ is about the word ‘cat’. I am not suggesting that metalinguistic facts are the real truth-makers for claims about essence, or any such thing. What I am suggesting is that despite obvious and important differences in content, claims about a word’s meaning and claims about a thing’s essence nevertheless share an important underlying pragmatic function—they allow us to make our constitutivity commitments explicit. The shared pragmatic function of these semantically distinct vocabularies is what explains the striking parallels, illustrated by Fine, between meaning and essence on the one hand, and analyticity and necessity on the other. Indeed, it is this underlying function which unites, at the pragmatic level, the family of concepts Quine characterizes in “Two Dogmas” as an “unbreakable intensional circle.”

3.4 Constitutivity Commitments and Intensional Vocabulary

The following is my account of what it is for an individual to treat a certain sentence as a constitutive principle in a particular conversational context:

Speaker-Relative Constitutivity: For interlocutor $i$ at time $t$, sentence $S$ counts as a constitutive principle governing expression $E$ just in case:

(1) $i$ has a practical commitment at $t$ to treat any (well-formed) $E$-utterance that implies the negation of $S$ as either:

(a) *prima facie* evidence of the utterer’s lack of competence with $E$; or

(b) *prima facie* evidence of the utterer’s meaning-divergence with respect to $E$; or

(c) *prima facie* evidence of a (non-empirical) disagreement with the utterer about the metaphysical nature of $E$’s subject matter.$^{144}$

---

$^{144}$ That is, the object to which $E$ refers or the objects that fall under $E$’s extension.
As Putnam’s examples from the history of science demonstrate, our entitlement to make the free moves normally licensed by our constitutivity commitments is not empirically indefeasible, and may in various cases be overridden. In some cases, however, our entitlement to the free moves is not taken to be sensitive to empirical evidence one way or another. Consider the basic axioms of arithmetic, or the introduction and elimination rules for the basic sentential connectives. Number terms like ‘2’ and sentential operators like ‘&’ are taken to be paradigm examples of the sorts of concepts whose meaning is established by implicit definition, and the sentences which play the role of implicit definers are just those sentences that we would expect competent speakers to treat as constitutive. Moreover, the sense of certainty that we have in the truth of such principles reflects a distinctive aspect of logico-mathematical practice: namely, that the kind of empirical mistake one can make while proving a theorem is importantly different from the kind of empirical mistake open to a biologist or physicist or chemist. In particular, the empirical mistakes possible in logic and mathematics (a) can be detected easily, without need of any special technology or measuring equipment; and (b) are not likely to occur systematically across the entire community.\(^\text{145}\)

It should be noted that interpreters have no choice but to solve for both meaning and belief simultaneously. If a person’s beliefs are too bizarre—that is, too divergent from those of the interpreter—it will be impossible for the interpreter to figure out what that person means by her utterances.\(^\text{146}\) In such cases, the salient question for the interpreter is no longer *Is my interlocutor speaking the same language as I am?* but rather *Is my interlocutor speaking a language at all?* Although deciding which of these questions is appropriate may appear

---

straightforward in familiar interpretive environments, there is no algorithmic procedure that yields a determinate verdict for all possible cases.

Clause (1) specifies the structure of the practical interpretive commitment—what I’m calling a constitutivity commitment—that one undertakes by virtue of treating a certain sentence as a constitutive principle. As practical commitments, constitutivity commitments are commitments to do something—viz., to interpret one’s interlocutors in a particular way. Each of the three sub-clauses—(1a), (1b) and (1c)—reflects an important cluster of interpretive dispositions typically exhibited by competent language-users. Indeed, part of what it is to be a competent language-user is to have these sorts of interpretive dispositions with respect to a suitable range of expressions. When we deploy intensional vocabulary (i.e., when we make meaning-claims or essence-claims), we thereby indicate that we have certain of these clusters of dispositions and that we expect the interpretive dispositions of other competent speakers to be sufficiently similar.

Sub-clause (1a) covers cases like the one discussed by Grice and Strawson in which a person puzzlingly claims that his neighbor’s three-year-old child is an adult. In their example, the speaker rejects any paraphrase he is offered (e.g., ‘You mean he is uncommonly sensible or very advanced for his age’, or ‘You mean he is already fully developed or won’t grow anymore’, etc.), and continues to insist that the child is literally an adult. Grice and Strawson note that in such a situation, it would be appropriate to “suspect that [the speaker] just does not know the meaning of some of the words he is using.”147 (And again, it should be noted that if the case is extreme enough, the interpreter may not be in a position to recognize the speaker’s behavior as genuinely linguistic.)

147 See Grice and Strawson (1956), pp. 204-205.
There may also be cases of incomplete competence. Consider, for example, Tyler Burge’s character who, despite having a good number of true beliefs about arthritis, takes himself to have arthritis in his thigh. Because Burge’s character is disposed to defer to experts and modify his usage upon learning of his mistaken belief about where arthritis can occur, it is plausible that he should not be counted as entirely incompetent with the term ‘arthritis’. True, his initial claim that he has arthritis in his thigh implies the negation of a sentence that we, and the relevant medical experts, are disposed to treat as a constitutive principle (e.g., ‘Arthritis is exclusively an ailment of the joints’). But this is only *prima facie* evidence of incompetence, and it is arguably overridden by the further information we have. Since we have been told that Burge’s character intended all along to be guided by the appropriate constitutivity commitments (i.e., those endorsed by the experts), we ought to say that he was partially or incompletely competent with the term ‘arthritis’ prior to having his usage corrected by the doctor. Upon being appropriately corrected, someone with deferential dispositions sufficient for incomplete competence can thereby go on to achieve full competence.

Sub-clause (1b) covers cases like the one illustrated by Quine’s reaction to the deviant logician who rejects principles like non-contradiction and *ex falso quodlibet*. According to Quine,

[Deviant logicians] think they are talking about negation, ‘~’, ‘not’, but surely the notation ceased to be recognizable as negation when they took to regarding some conjunctions of the form \( p \cdot \sim p \) as true, and stopped regarding such sentences as implying all others. Here, evidently, is the deviant logician’s predicament: when he tries to deny the doctrine he only changes the subject.\(^{149}\)

\(^{148}\) See Burge (1979), pp. 104-107.
\(^{149}\) Quine (1986) p. 81.
This is not the sort of remark one would expect to hear from a philosopher who rejects any kind of distinction between constitutive and non-constitutive sentences, but Quine clearly needs some such distinction if passages like this are to be coherent.\textsuperscript{150} Quine’s most fundamental and compelling arguments for rejecting any such distinction derive from his skepticism about the explanatory role of meaning. My account of constitutivity commitments, however, should be perfectly acceptable by Quinean naturalist lights. For on my view, the primary function of intensional discourse (including constitutivity-talk) is not to report on explanatorily ineliminable meanings or essences, but to facilitate discursive coordination by allowing interlocutors to give voice to—and thus scrutinize, challenge, endorse, or commend—the interpretive dispositions of the various conversational participants. The reality that prompts us to engage in meaning-talk is a thinner one than classical realists about meaning have supposed, but this in no way undermines the legitimacy or utility of the vocabulary. For once we have given an alternative explanation of what we are doing when we employ such vocabulary, there is no longer any reason to feel metaphysical guilt about using it in ordinary, non-theoretical contexts.

Sub-clause (1c) covers cases in which a clash of interpretive dispositions manifests itself as an object-level (non-metalinguistic) disagreement that cannot be settled empirically. This includes typical disputes over ontology and essence, and also disputes about metaphysical necessity more generally.\textsuperscript{151} These are cases in which the interlocutors fight

\textsuperscript{150} See Burgess (2004). See also Arnold and Shapiro (2007).

\textsuperscript{151} See Thomasson (2007) for a similar suggestion. Her view, which she calls Modal Normativism, “denies that basic modal claims of metaphysical necessity are descriptive claims that need truthmakers at all, holding instead that claims of metaphysical necessity primarily serve the prescriptive function of expressing semantic rules for the terms used in them, or their consequences, by remaining in the object language” (p. 136). I agree with Thomasson that such modal claims do not require modal facts as truthmakers—their truth or falsity is not to be understood in terms of whether or not they “track” or “correspond to” objective modal features of the world. However, I would resist her implicit suggestion that to be a descriptive claim is to be in need of a truth-
with one another about what some object essentially is or what it must be if it is to exist at all. They are cases in which the disputing parties take themselves to hold incompatible theories of the nature of the subject matter in question, despite having access to all the same sensory information.

Although some basic set of empirical background beliefs is required before metaphysical inquiry can get off the ground, the answers to metaphysical questions cannot simply be read off of the empirical data. Instead, metaphysical inquiry proceeds primarily by tracing out what one takes to be the logical relations between certain clusters of beliefs. These inferential connections are typically assumed to be grounded in the nature of the subject matter those beliefs are about. In tracing out such connections, the metaphysician may feel that her job is to strip away the contingent, the historical, the interest-relative, and thereby make profound discoveries about the objective nature of things and come in this way to understand the demands that non-human reality makes upon our classificatory practices. But this Platonistic picture is rooted in a mistake; for only beings with interests, plans, or goals are capable of issuing or abiding by demands. Metaphysical discourse may nonetheless be useful, however, insofar as it allows us to make explicit and critically discuss the demands we place on ourselves and others.

It should be emphasized that it does not follow from the fact that sentence $S$ is being treated as a constitutive principle in context $C$ that sentence $S$ expresses a necessary truth in all contexts. It could turn out, as it did in the case of ‘kinetic energy’, that a sentence once taken to be constitutive is later shown to be empirically false. Yet there is an indirect connection between constitutivity and metaphysical necessity. In any context in which one

---

maker, for it seems to me that whether or not a claim counts as descriptive is best understood as grammatical issue, not a metaphysical one. For more on this, see Price (1997), (2004), (2009).
sincerely believes sentence $S$ and is entitled to treat $S$ as a constitutive principle, one is (defeasibly) entitled to assume that what $S$ says admits of no alternatives, i.e., that it is metaphysically necessary. If, however, one is presented with information that shows alternatives could be admitted, one’s entitlement is thereby compromised.

It is also important to note that what is indicated by the making of essence-claims and meaning-claims is not the same as what is said in making such claims. These are not autobiographical reports or claims about personal preference. Nor are the relevant facts about the speaker’s dispositions and expectations to be construed as the truth-makers for such claims. Like the expressivist metaethical views developed by Simon Blackburn and Allan Gibbard, the account I give here is intended to be naturalistic, but non-reductive. Unlike Blackburn and Gibbard, however, I do not think that the content of normative discourse (which includes intensional vocabulary on my view) is fundamentally different in kind from the content of non-normative discourse. Both Gibbard and Blackburn are committed to a mentalistic semantics according to which the content of normative claims is derived from desire-like (or planning) states of mind, while the content of non-normative claims is derived from belief-like or representational states of mind. Instead, following Huw Price, I hold that the difference between normative and non-normative discourse is to be understood as a difference of pragmatic function rather than a difference of semantic content. Against the backdrop of a thoroughgoing deflationism or minimalism about semantic notions like truth, reference, facts, and propositions, it would be a mistake to characterize expressivism in semantic terms. For qua deflationist, I deny that semantic notions are fit to do the requisite explanatory work.

---

Formal semantics is, of course, an important and useful project. But formal semantic theory is in the business of codifying competent language use by constructing a formal recursive system that models the infinitely productive capabilities of language-users. My project, by contrast, is that of explaining why humans developed intensional vocabulary in the first place by describing its pragmatic function. As I conceive them, these two projects are not rivals.\textsuperscript{154}

3.5 Constitutivity, Retrievability, and Formal Pragmatics

In this section, I draw upon the resources of formal scorekeeping pragmatics in order to provide an account of what it is for a group of interlocutors to share constitutivity commitments in a discourse interaction. To this end, I outline Craig Roberts’s formal theory of context of utterance, and argue that it should be modified to include parameters for the interlocutors’ constitutivity commitments.\textsuperscript{155}

Following Lewis, Roberts conceives of context on the model of a dynamic scoreboard on which information about the current state of the discourse interaction is recorded.\textsuperscript{156} The scoreboard is dynamic in the sense that it gets updated whenever score-affecting moves are made by the interlocutors. Following Grice, she takes discourse to be a rational enterprise organized around conversational goals and the strategies that conversational participants deploy to achieve those goals.\textsuperscript{157} Following Stalnaker, she holds that the primary or overarching goal in conversation is that of sharing information about the

\textsuperscript{154} In this, I agree with Price (2004): “A modest Fregean content-specifying theory is compatible with the view that the interesting theoretical vocabulary for linguistic theory is pragmatic or use-theoretic, rather than semantic” (p. 215).


\textsuperscript{156} See Lewis (1979).

\textsuperscript{157} See Grice (1957), (1967).
The shared information about the world in any particular conversation is the *common ground*, i.e., the set of propositions that all participants currently treat as true. According to Roberts, the goal of sharing information about the world can be thought of as an intention to answer the Big Question: *What is the way things are?* But in order to do this, interlocutors have to develop more manageable sub-goals, which can be represented by a series of sub-questions. Each sub-question reflects the current topic of the conversation, and a sub-question is only removed from the questions-under-discussion stack when a complete answer to it is entailed by the common ground.

In addition to the *discourse goals* that correspond to sub-questions derived from the Big Question, interlocutors also have *domain goals*—i.e., “deontic priorities [that] generally direct the type of inquiry which we conduct in conversation, the way we approach the question of how things are.” Discourse goals thus reflect our *theoretical* intentions while domain goals reflect our *practical* intentions. And, as Roberts emphasizes, our practical (domain) goals direct and constrain the strategies we employ in service of our theoretical (discourse) goals. Roberts gives a nice illustration of this phenomenon:

For example, if a hostile witness in a trial is asked whether the accused took money from his firm and answers *He regularly wrote checks to cover his expenses*, one should not take this reply to necessarily be a complete answer to the question, as all good prosecuting attorneys know. That is, the witness will uncooperatively construe the question itself in the narrowest possible way, to avoid having to lie or give information that might serve the prosecution; but given her obvious overarching personal goal of giving as little information as possible, this is predictable.

---

159 Each time a new proposition is entered into the common ground, alternative ways the world might be are ruled out, thereby shrinking what Stalnaker calls the *context set*—i.e., the set of possible worlds in which all the propositions in the common ground are true. It should be noted that the information in the common ground reflects not only the explicit assertions accepted by the discourse participants, but also represents shared background information, including perceptual information about salient objects in the interlocutors' vicinity, etc.
161 Roberts (2012), p. 5.
In this example, one of the hostile witness’s domain goals—in particular, her desire to avoid incriminating the accused—shapes and constrains the quantity and quality of the information she provides to her interlocutor.

If a discourse participant is sincere, competent, and cooperative, she must also be committed to a practical meta-goal regarding the utterances she produces over the course of a conversation. In particular, whenever she makes an utterance, she must intend (a) that her utterance will have a certain effect on her audience, (b) that her audience will recognize her intention to produce that effect in them, and (c) that her utterance will have the intended effect because her audience has recognized her intentions. These Gricean considerations entail what Roberts calls the principle of Retrievability:

**Retrievability:** In order for an utterance to be a rational, cooperative act in a discourse interaction $D$, it must be reasonable for the speaker to expect that the addressee can grasp the speaker's intended meaning in so-uttering in $D$.\footnote{Roberts (2012), p. 3.}

I want to suggest that the overarching practical meta-goal embodied in the principle of Retrievability is achieved in part by the undertaking of constitutivity commitments. Indeed, the undertaking of constitutivity commitments is an essential aspect of what it is to be a competent, cooperative discourse participant because sufficient interpretive alignment among the interlocutors is a precondition on the possibility of information sharing in any given context.

On my proposal, constitutivity commitments are generated as a result of each interlocutor's meta-commitment to respect Retrievability by using shared expressions properly or correctly. According to Roberts, “the structure of a discourse interaction is designed to help

---

162 See Grice (1957).
163 Roberts (2012), p. 3.
satisfy Retrievability."\textsuperscript{164} Since the point of undertaking constitutivity commitments is also to help satisfy Retrievability, my proposal to expand the conversational scoreboard to include parameters for the constitutivity commitments of the interlocutors is well-motivated.

Just as interlocutors can provide each other with information about their domain goals, they can also provide each other with information about their constitutivity commitments—which is precisely what they must do when communication breaks down and there is a need for interpretive coordination. Recall that discourse goals reflect our desire for information about the world, while domain goals are “goals in the real world, things we want to achieve quite apart from inquiry” and it is these latter, practical goals that shape and constrain how we talk and think about the world.\textsuperscript{165} Like domain goals, constitutivity commitments embody practical, rather than theoretical, intentions. As such, the scoreboard parameters for individual and shared constitutivity commitments can be represented in a way that is formally analogous to the way that individual and shared domain goals are modeled. (See the appendix for the technical details.)

The most straightforward way to ensure that one is using shared expressions correctly is to use them in (what one takes to be) their conventional or canonical senses. But with proper stage-setting, interlocutors can institute alternative standards of correctness in particular contexts, thereby forging new linguistic tools which are better suited to their common goals. This standard-shifting can be explained in terms of the interlocutors’ collaborative negotiation of shared constitutivity commitments. By strategically calibrating these commitments, discourse participants are able to extend and enrich existing vocabulary,

\textsuperscript{164} Ibid., p.4.
\textsuperscript{165} Roberts (1996), p. 5.
allowing them to pursue goals (practical and theoretical) that would have been difficult or impossible to achieve otherwise.

3.6 Kripkean Essentialism: A Neglected Alternative?

With the formal apparatus described in the previous section, we can model what is going on when interlocutors assess each other for linguistic competence or negotiate the meanings of shared expressions in a discursive context. In the same way, we can also model what is going on when interlocutors engage in disputes over the “very being of a thing, whereby it is what it is”—i.e., disputes over the correct identity, existence, and persistence conditions of various objects.

One potential worry about my account is that I have been discussing these issues in terms of Kit Fine’s “neo-Aristotelian” views about essence, and in so doing I have ignored another prominent form of essentialism that derives from suggestions made by Kripke and Putnam.166 That is, in addition to Finean Essentialism, there is also Kripkean Essentialism, on which essence is conceived in importantly different terms. According to the Kripkean alternative, the essential properties of an object are those microstructural properties to be discovered \textit{a posteriori} by our best sciences. This sort of view identifies essence with something scientifically respectable—viz., the microstructural properties that explain the salient surface features of the thing in question—and renders knowledge of essence \textit{a posteriori} rather than \textit{a priori}. Thus my deflationary view that essences play no explanatory role seems to be threatened by the availability of the Kripkean alternative. For if essential properties just are those properties that our best sciences are built to track, then such

\footnote{166 See Kripke (1980) and Putnam (1975).}
properties would seem to be genuinely explanatory, and the truth-makers for essence-claims would seem to be empirical facts about real microstructural properties.

It is not clear, however, that one can coherently endorse Kripkean Essentialism without also endorsing a number of other controversial theses about language and meaning. Kripkean Essentialism is typically held as part of a theoretical package according to which names and natural kind terms are rigid designators with no content over and above their referents, and the reference of such terms is fixed by baptismal episodes in which we stipulate that we intend the term to pick out a certain individual or kind with which we are acquainted.\footnote{This characterization in terms of explicit stipulation is idealized. The basic point is not that all such terms were actually introduced in this way, but that they behave \textit{as if} they were stipulated in this manner. What really matters to proponents of this view is that when people use such a term, they do so with the intention that it is to apply to all instances of the relevant kind. Using terms with such an intention is supposed to be sufficient for the reference-fixing to occur. See Soames (2007) for further discussion.}

But even if we suppose that Kripkean Essentialism can be coherently endorsed without also buying into this (by my lights, dubious) background theory, fundamental difficulties remain. In particular, the reduction of essence to microstructure either fails to accommodate the actual use of essence-talk, or else constitutes an implausibly revisionist account of such talk. Acceptance of the reductive move forces one to hold that only things that have scientifically detectable microstructural properties—perhaps water or whales or the color green—have essences. But this move undermines the legitimacy of making essence-claims about things like democracy or personhood or socialism or marriage. The Kripkean must hold that such things lack essences, since they clearly lack the relevant sort of microstructure, and that there is thus something systematically wrong with large swaths of what otherwise looked like competent use of the term ‘essence’.
Such considerations greatly undermine the plausibility of Kripkean Essentialism. My view, by contrast, begins with the observation that competent speakers make essence-claims about things with, and things without, scientifically discoverable microstructures. It is not clear that there is an identifiable microstructure that unifies all the things we call planets, and there certainly is no such microstructure underlying bachelorhood, yet interpretive coordination may well be facilitated by claims about the essential characteristics of planets or bachelors. The same points apply to essence-claims about fictional or mythical things like witches or unicorns or deities. I explain this data via a unifying pragmatic account of what we are doing when we make essence-claims (expressing constitutivity commitments) and why essence-talk is useful to us (it facilitates interpretive coordination).

There is, however, an important insight to be gleaned from Kripkean Essentialism regarding the constitutive principles associated with terms that we intuitively think of as picking out so-called “natural kinds.” In particular, the insight is that if a certain expression is a natural kind term, then anything that falls under its extension must have a certain scientifically discoverable microstructure which explains the superficial features of the kind, even though we may not yet know what that microstructure is. This suggests that (alleged) natural kind terms are to be governed by schematic constitutive principles of roughly the following form: If expression $E$ is a natural kind term, then there is some scientifically discoverable microstructural property $P$, such that if $E$ correctly applies to some object, then that object must have $P$.

Kripkeans are quite right to insist that our knowledge of these microstructural properties is \textit{a posteriori} rather than \textit{a priori}. Nevertheless, on the assumption that $E$ is a natural kind term, it follows immediately that whatever $E$ denotes has a certain microstructural property $P$ necessarily, even if we don’t currently know what property that is. In this way, our
entitlement to treat ‘Water is H₂O’ as a necessary truth ultimately flows from our constitutivity commitments for the term ‘water’.\textsuperscript{168}

There are a great many ways to describe reality, but the kind of description (as well as the vocabulary) that is appropriate in a given context will depend on what the interlocutors care about in that context. Since the goals that typically underwrite scientific inquiry arise out of a desire to predict and explain natural phenomena, the vocabulary that we develop for this purpose ought to be governed by constitutive principles which help facilitate prediction and causal explanation. When certain principles fail to facilitate those aims, we have grounds for revoking their status as constitutive. But it is important to remember that our goals may be quite different from context to context, and this will in turn affect what we take to be constitutive.

3.7   Concluding Remarks

Our human tendency to conform to the usage-patterns of those we recognize as competent speakers makes communication easier and more efficient by reducing the need for constant interpretive re-calibration. I suspect that this conformism also underlies the naïve intuition that our expressions (or at least most of them) have determinate application conditions. In particular, when certain clusters of constitutivity commitments become sufficiently well-entrenched—that is, when they are widely and persistently acknowledged across many different conversations by many different speakers—certain expressions will appear to have fixed application conditions which uniquely determine what counts as correct usage in all possible situations.

\textsuperscript{168} See Thomasson (2007) for a similar suggestion couched in terms of schematic linguistic rules.
For those of us impressed by Wittgensteinian rule-following considerations, however, there is good reason to believe that linguistic proprieties cannot be grounded in facts about our usage-dispositions or semantic intentions.\textsuperscript{169} What is needed instead is a picture of linguistic activity on which regularities of use emerge from and evolve within an ongoing process of interpretive negotiation. Sellarsians have long emphasized that the functional role of linguistic expressions includes ties to perception (language-entry transitions), ties to action (language-exit transitions), and inferential structure (language-language transitions).\textsuperscript{170} I submit that shared constitutivity commitments determine, in any given discourse interaction, which language-language transitions are to be counted as free moves in the inferential game. Rational negotiation of these commitments is a crucial aspect of human linguistic practice without which information-sharing and cooperative inquiry would be impossible.

\textsuperscript{169} See Wittgenstein (1953) §§138-242. See also Kripke (1982), and Wilson (1994).
Chapter 4
Constitutivity and Quine’s Challenge

4.1 Introduction

From the perspective of the competent language-user, constitutive principles say what usually goes without saying. When such sentences are employed, the primary point of doing so is to call attention to a standard for the proper application of a word rather than to convey interesting or relevant information about the world. For example, I typically don’t go out of my way to utter ‘If X is warmer than Y, and Y is warmer than Z, then X is warmer than Z’ unless I suspect my interlocutor to be misusing or misinterpreting the phrase ‘warmer than’.

It is standardly assumed that the viability of any sort of distinction between constitutive and non-constitutive uses of linguistic expressions depends on the viability of the analytic-synthetic distinction, but this assumption is mistaken. Competent interpreters distinguish between what they take to be constitutive and non-constitutive sentences not because they are tracking genuine analyticity but because doing so facilitates interpretive coordination among interlocutors with potentially divergent perspectives. One can reject the analytic-synthetic distinction on Quinean grounds and still coherently endorse a distinction between changes of belief and changes of meaning. In order to be entitled to do this, however, one must answer a fundamental challenge Quine raises against Carnap, but which
Carnap refuses to take up: viz., one must identify some sort of behavioral criteria by means of which the constitutive status of a sentence can be determined. In this chapter, I argue that my account of interpretive coordination provides a satisfactory answer to this challenge, and does so in a way that Quine himself ought to accept. I also argue that if we are to adequately explain the normative dimension of discursive practice, we require a criterion like the one my account provides—namely, one which is characterized in terms of behavior more sophisticated than mere assent.

4.2 Why Quine Needs an Account of Constitutivity

In light of his rejection of analyticity, we should expect Quine to avoid any mode of argumentation that looks like conceptual analysis or that otherwise implies endorsement of a distinction between matters of meaning and matters of fact. As such, it is puzzling to come across passages such as this one from *Word & Object*:

> To take the extreme case, let us suppose that certain natives are said to accept as true certain sentences translatable in the form ‘p and not p’. Now this claim is absurd under our semantic criteria. And, not to be dogmatic about them, what criteria might one prefer? Wanton translation can make natives sound as queer as one pleases. Better translation imposes our logic upon them, and would beg the question of prelogicality, if there were a question to beg.

Here we have Quine invoking “semantic criteria” and insisting that good translation requires reading our own (i.e. classical) logic into our interlocutor’s language. He claims that he doesn’t want to be dogmatic about these criteria, but he clearly takes himself to be entitled to make arguments based on them. Quine seems to be saying that a charitable assumption of rationality is constitutive of the notion of *good translation*. Does this mean that Quine thinks

---

171 See Arnold and Shapiro (2007).
172 Quine (1960), p. 58, note omitted.
that sentences of the form ‘X is a good translation only if X proceeds according to a
charitable assumption of rationality’ are true by virtue of meaning? Perhaps not, if analytic
truth is the kind of thing one is supposed to be dogmatic about.

Interpretive misalignment is a chronic cause of communication breakdown. In the
face of such a breakdown, it is often useful to ask the following sort of question: Does my
interlocutor have odd beliefs about Fs, or does she just mean something different from what
I mean by the predicate ‘F’? However, in the absence of a viable analytic-synthetic
distinction, there is supposed to be no way to adjudicate this kind of question, no fact of the
matter. Curious, then, that Quine feels entitled to argue as follows:

[Deviant logicians] think they are talking about negation, ‘¬’, ‘not’; but surely the notation
cessated to be recognizable as negation when they took to regarding some conjunctions of the
form ‘p · ¬p’ as true, and stopped regarding such sentences as implying all others. Here,
evidently, is the deviant logician’s predicament: when he tries to deny the doctrine he only
changes the subject.173

Quine is saying, quite explicitly, that if you think there are true contradictions, then you don’t
mean negation when you use the word ‘not’—you have simply changed the subject. But this
could only amount to a change of subject if the shape of the subject matter is somehow
constituted (at least partially) by a principle that the deviant logician denies, i.e., the law of non-
contradiction. Once again, it looks as if Quine has admitted that it is possible to usefully
distinguish a change of meaning from a change of belief—at least when it comes to logic.

Donald Davidson, too, engages in this style of argumentation, despite explicitly
endorsing Quine’s rejection of analyticity.174 In fact, when Davidson argues this way, he
often does so with less hesitation and hedging than does Quine. In the first passage quoted
above, Quine qualifies his appeal to semantic criteria by saying that he doesn’t want to be

dogmatic about them, which suggests that he thinks it is at least possible that the criteria he has in mind could turn out to be wrong. Davidson, by contrast, is often much bolder:

The methodological advice to interpret in a way that optimizes agreement should not be conceived as resting on a charitable assumption about human intelligence that might turn out to be false. If we cannot find a way to interpret the utterances and other behavior of a creature as revealing a set of beliefs largely consistent and true by our own standards, we have no reason to count that creature as rational, as having beliefs, or as saying anything. Here we have Davidson saying that nothing counts as a rational, language-using entity unless it has mostly true beliefs. If this is right, then anyone who insists that it is in principle possible for a rational creature to fail to have mostly true beliefs is either talking nonsense or else has attached a different meaning to the phrase ‘rational creature’. Indeed, Davidson thinks that a whole family of notions—including truth, rationality, and belief—are intricately and inextricably bound up with one another in this way. Thus Davidson, like Quine, routinely presupposes a distinction between matters of meaning and matters of belief.

A case can be made that Quine’s so-called “Circularity Argument” from “Two Dogmas” makes no sense unless some such distinction is presupposed. The basic line of that argument goes as follows: The term ‘analytic’ belongs to a family of intensional expressions—including ‘synonymous’, ‘necessary’, ‘definition’, and ‘semantical rule’—and since there is no way to say what analyticity essentially is except by appeal to one or more of these related notions, there is no way to break out of the intensional circle and thus no way to render the analytic-synthetic distinction naturalistically acceptable. Note, however, that the very idea of an “intensional circle” of this kind is simply incoherent unless we assume that at least some of the inferential connections between these notions have constitutive status. For to conceive of this set of concepts as a tightly-knit intensional family just is to hold that

175 Davidson (1973), p. 137.
certain principles linking these concepts cannot be denied without changing the subject. If one looks closely, one can find several of these constitutive linking principles explicitly formulated in the text of “Two Dogmas” itself. For example, Quine explicitly states that the adverb ‘necessarily’ is to be “so construed as to yield truth when and only when applied to an analytic statement.”¹⁷⁶ A few pages later, he tells us that “[s]tatements may be said to be cognitively synonymous when their biconditional (the result of joining them by ‘if and only if’) is analytic.”¹⁷⁷ If the suggestion that analyticity, synonymy, and necessity are members of an unbreakable intensional circle is to make any sense, Quine must treat these claims as constitutive principles that serve to bind the family together.

Later on, in *Word & Object*, Quine pushes this line of argument even further by attempting to construct behavioristic replacements—e.g., “stimulus meaning,” “stimulus-synonymy,” and “stimulus-analyticity”—which he defines in terms of the sensory stimulations that would prompt a speaker to assent to or dissent from a queried sentence in various circumstances.¹⁷⁸ But according to Quine, it is easy to see that “stimulus meaning as defined falls short in various ways of one’s intuitive demands on ‘meaning’ as undefined, and that sameness of stimulus meaning is too strict a relation to expect between a native occasion sentence and its translation—even in so benign a case as ‘Gavagai’ and ‘Rabbit’.”¹⁷⁹ Here, we have Quine saying that his notion of stimulus meaning “falls short” of being an adequate replacement for the ordinary notion of meaning because it fails to meet the “intuitive demands” that he associates with the latter. Since the “intuitive demands” are not

¹⁷⁶ Quine (1951), pp. 29-30.
¹⁷⁷ Ibid., p. 32 note omitted.
¹⁷⁸ Quine (1960), chapter 2.
¹⁷⁹ Ibid., p. 39.
met in this case, Quine in effect concludes that to explicate intensional notions in terms of stimulus meaning is to change the subject.180

Along similar lines, John Burgess has argued that Quine cannot explain the felt obviousness of elementary mathematics unless he endorses a distinction between sentences that are part of the meaning of an expression and those that are not. Quine criticizes the so-called “linguistic doctrine of logical truth” on the grounds that to say that logic is true by convention is no more explanatory than to say that it is obvious. But he has a hard time accounting for the fact that ‘2+2=4’ feels intuitive and self-evident in a way that ‘There are brick houses on Elm Street’ does not.

Complex mathematical propositions are not always intuitively or self-evidently true, but one can be confident in accepting such a proposition provided that one can derive it on the basis of a series of obvious steps. In order to make sense of such phenomena, Burgess argues that Quine needs to countenance a distinction between the “basic” principles, which present themselves as obvious, and the “non-basic” ones, which do not. Burgess thus offers the following suggestion:

My proposal is that the law should be regarded as ‘basic’, as ‘part of the meaning or concept attached to the term’, when in case of disagreement over the law, it would be helpful for the minority or perhaps even both sides to stop using the term, or at least to attach some distinguishing modifier to it. Such basic statements would then count as analytic, as would their logical consequences, at least in contexts where, in contrast with the examples above, there is no disagreement over logic. This proposal makes the notion of analyticity vague, a matter of degree, and relative to interests and purposes: just as vague, just as much a matter of degree, and just as relative to interests and purposes as ‘helpful’. But the notion, if vague, and a matter of degree, and relative, is also pragmatic, and certainly involves no positing of unobservable psychological entities, and for those reasons seems within the bounds of what a Quinean could accept.181

180 This, of course, he is happy to do when it comes to linguistic theorizing. But this is not exactly eliminativism. As Robert Kraut has reminded me, Quine has no qualms with the ordinary use of intensional vocabulary—including talk of necessity and meaning—and is happy to recognize the pragmatic utility of such talk.

The idea is that when there is a disagreement over a sentence taken to be “basic” in Burgess’s sense, it will be helpful for one or both parties to rephrase things in other terms. What Burgess is gesturing at in this passage is just the sort of thing that my account of interpretive coordination via the alignment of constitutivity commitments is intended to capture. As it stands, however, the proposal Burgess sketches is not only underdeveloped but crucially incomplete. What is needed, and what my account provides, is an answer to Quine’s challenge—i.e., a specification of some basic behavioral criteria by which to ascertain the constitutive status of various sentences in various contexts.

4.3 Quine’s Challenge: Providing Behavioral Criteria

Before considering Quine’s challenge directly, let us recall Carnap’s mature views on meaning and analyticity. Following Kant, Carnap holds that the mathematical and logical elements in our scientific theories give fundamental structure to scientific inquiry, in the sense that these logico-mathematical elements provide the inferential and evidential frameworks that make empirical testing and experimentation possible.\(^\text{182}\) Without the frameworks provided by logic and mathematics, crucial notions like confirmation and evidence would be unintelligible and science itself would be impossible. The idea is that some such logical framework must be in place, or else we could have no grasp of evidential or inferential relationships between propositions.

Unlike Kant, however, Carnap is a pluralist about these structure-giving frameworks. He rejects the idea that there is one “true” or “correct” or “necessary” framework for inquiry, insisting instead that theorists should adopt an attitude of tolerance toward the various

\(^\text{182}\) For further details and historical context, see Friedman (2006), (2008), and (2012). See also Creath (2007).
proposals, and assess them on the basis of their *utility* rather than their objective correctness.\(^{183}\) For example, intuitionists and classical logicians should not see their dispute as being about which one of them has accurately described the “correct” logic. Instead, proposals regarding the logico-mathematical frameworks at work in science should be assessed on the basis of their utility relative to our current scientific goals. Thus, Carnap advocates the so-called Principle of Tolerance: As theorists, we are free to choose whichever framework we deem most useful for our purposes—our choice of logic is constrained only by our own theoretical goals, whatever those may be.

Time and again, Carnap laments the fact that progress in human inquiry (including progress in science and philosophy) is often frustrated because rival theorists tend to talk past one another. Indeed, a deep motivation behind Carnap’s views on analyticity and tolerance is his desire for clear, efficient communication between theorists. In his “Intellectual Autobiography,” he writes:

> Even in the pre-Vienna period, most of the controversies in traditional metaphysics appeared to me sterile and useless. When I compared this kind of argumentation with investigations and discussions in empirical science or in the logical analysis of language, I was often struck by the vagueness of the concepts used and by the inconclusive nature of the arguments. I was depressed by disputations in which the opponents talked at cross purposes; there seemed hardly any chance of mutual understanding, let alone agreement, because there was not even a common criterion for deciding the controversy.\(^{184}\)

Carnap’s point is that communication between a pair of interlocutors on a particular topic will fail unless they are roughly in agreement on the evidential and inferential standards that they take to be in play with respect to the topic at hand. Such failures can be avoided, Carnap suggests, by explicitly and precisely articulating the evidential and inferential standards which are to govern some chunk of vocabulary. With those standards laid bare,

\(^{183}\) See Carnap (1934/1937), p. 52. See also Carnap (1950), p. 221.
\(^{184}\) Carnap (1963), pp. 44-45.
theorists are in a much better position to evaluate the costs and benefits of adopting the proposed theory and its attendant vocabulary.

Carnap sees the analytic-synthetic distinction as a piece of technical machinery specifically designed to allow theorists to precisely articulate the structure of various frameworks so that those frameworks can be clearly understood, compared, and assessed. For Carnap, characterizing a sentence as analytic is not to make a descriptive empirical claim about how that sentence-type is actually used in natural language. Rather, the function of an analyticity attribution is to express an explicit linguistic proposal in the context of a particular project of conceptual engineering or language-planning. Armed with these tools, the scientific-minded philosopher is free to abandon traditional metaphysics and epistemology, and engage instead in a project of linguistic engineering Carnap calls *the logic of science*. The task of the logician of science is to construct rigorous explications of various linguistic frameworks and offer pragmatic assessments of their utility for various scientific purposes.

It should be noted that for Carnap, the adoption of a particular framework is never a permanent or unrevisable commitment. This point is crucial, and often overlooked. If ever a framework becomes cumbersome or inadequate to our needs, it may be modified or abandoned in favor of a more useful framework. For this reason, it is misleading to suggest, as Quine does, that Carnap takes analytic sentences to be true “come what may.”\(^{185}\) So long as they are playing the role of semantical rules, it does not make sense to seek evidence for or against them, since the job of such rules is precisely to specify what counts as evidence. However, their status as *rules* may change—that is, we may choose to stop regarding them as framework constituting principles, at which point it becomes possible to reject them on

---

\(^{185}\) Quine (1951), p. 43.
empirical grounds.\textsuperscript{186} One important upshot of this is that for Carnap, the analytic-synthetic distinction is not absolute, but is instead always relative to a particular framework. As a result, Carnap’s views turn out to be consistent with Duhemian epistemological holism.\textsuperscript{187}

Against the backdrop of Carnap’s language-planning project, let us now consider Quine’s challenge. Richard Creath gives perhaps the clearest and most plausible description of just what Quine expects Carnap to provide. Creath articulates the challenge as follows:

So what does Quine want? In “Two Dogmas” there is at most a hint that we might not even notice but for the rest of his writings. Both before “Two Dogmas” and repeatedly after it Quine insists that he must be provided “behavioral criteria” for all intelligible terms and for “analytic” in particular. What Quine is demanding for analyticity, then, is essentially what Carnap demanded for physical length and other notions that were suitable for empirical science, namely, an indication in observational terms as to when the use of these notions was appropriate. This might fall short of a complete definition and the link to the observable might be indirect, but that link had to be there. What Quine takes to be observable, in the case of language, is behavior, and hence the expression, “behavioral criteria.”\textsuperscript{188}

Creath suggests that Carnap thinks of his project as akin to the work of the pure geometer, and thus rejects Quine’s demand for behavioral criteria as inappropriate. For Carnap, analytic sentences do not make empirical claims about the world but rather constitute the meanings of expressions in a formalized system, thereby giving fundamental structure to the language. Because he thinks of himself as engaged not in empirical linguistics but in a project of post-Kantian metamathematics, Carnap concludes that Quine’s challenge simply misses the point.

Carnap is surely right to insist that there must be an intellectual division of labor which allows room for specialists to focus on purely formal matters (e.g., pure geometry as

\textsuperscript{186} See Carnap and Quine (1990), where Carnap says, “naturally the same sentence (i.e., the same sequence of words or symbols) can be analytic in one system and synthetic in another, which replaces the first at some time” (p. 432).

\textsuperscript{187} See Carnap (1934/1937), where he says that “there always exists the possibility of maintaining the hypothesis and renouncing acknowledgement of the protocol-sentences” and concludes that an empirical “test applies, at bottom, not to a single hypothesis but to the whole system of physics as a system of hypotheses (Duhem, Poincaré)” (p. 318).

\textsuperscript{188} Creath (2007), p. 328.
opposed to applied geometry). However, there may be more to Quine’s challenge than Carnap thinks. Here is Creath:

It is all very well to say that Carnap is making proposals rather than describing natural languages, but he is not even making proposals unless we know what it would be for someone to accept and act on them. Nor could we evaluate them for utility unless we know what it is to adopt such proposals. Perhaps continuing the analogy with geometry will make this clearer. As Quine pointed out many years before “Two Dogmas,” a pure mathematical geometry completely divorced from methods of linear measurement can be thought of as a mathematical rather than an empirical theory (Quine, 1936). But the cost of this is that it ceases to be a theory specifically about space; it would no more be about spatial points than it would be about ordered n-tuples of real numbers or for that matter about any sufficiently numerous domain whatsoever. Only when some methods of measurement are supposed as waiting in the wings does the abstract formalism earn its claim to be a geometry at all. So it is with Carnap’s metamathematics of language. Without appropriate linkages to behavior, even proposals are no more than abstract formalisms that forfeit their claims to be about language at all.180

The point is that if the notion of analyticity is to have anything to do with actual human language, we need some account of how to tell the difference between an analytic proposal and a substantive descriptive claim. That is, we need to know what kind of evidence to look for in a person’s behavior that would indicate that she holds a certain sentence to be analytic as opposed to synthetic. Until some such account is provided, we have no right to consider Carnap’s linguistic frameworks to be anything more than abstract formalisms. Rendered in this way, it appears that Quine’s demand for behavioral criteria is not entirely inappropriate after all.

4.4 Assent Dispositions and Language Learning

To accept Quine’s challenge is to agree only that some sort of behavioral criteria must be specified, but according to Creath, it does not follow that the account needs to be particularly sophisticated. As Creath points out, this concession to Quine is actually quite

180 Ibid., pp. 330-331.
narrow. All it would take to meet this modest challenge would be to specify crude or preliminary criteria, which, Creath claims, is not particularly difficult to do. In fact, by Creath’s lights, Quine himself provides such criteria in later work produced after Carnap’s death. In particular, Creath mentions the account of analyticity sketched briefly in *Roots of Reference*. Here is the relevant passage from Quine:

In *Word and Object* I defined a *stimulus-analytic* sentence as one to which every speaker is disposed to assent. The analytic sentences in the present sense are a subclass of those, and a somewhat nearer approximation to the analytic sentences uncritically so called. Even so, we have here no such radical cleavage between analytic and synthetic sentences as was called for by Carnap and other epistemologists. In learning our language each of us learns to count certain sentences, outright, as true; there are sentences whose truth is learned in that way by many of us, and there are sentences whose truth is learned in that way by none of us. The former sentences are more *nearly* analytic than the latter. The *analytic* sentences are the ones whose truth is learned in that way by all of us; and these extreme cases do not differ notably from their neighbors, nor can we always say which ones they are.190

We are thus given the following suggestion: The degree to which a sentence may be called analytic depends on what percentage of the community’s members learned to count it true in the process of learning language. This yields a behavioral characterization of analyticity in terms of *community-wide, language-learning-based dispositions to assent to a sentence*.

Years later, in “Two Dogmas in Retrospect,” Quine mentions his *Roots of Reference* account approvingly, and then proceeds to reformulate it in importantly different terms. He cheerfully proclaims that “[a]nalyticity undeniably has a place at a common-sense level,” and that “[i]t is intelligible and often useful in discussion to point out that some disagreement is purely a matter of words rather than of fact.”191 Then he writes the following:

In *Roots of Reference* I proposed a rough theoretical definition of analyticity to fit these familiar sorts of cases. A sentence is analytic for a native speaker, I suggested, if he learned the truth of the sentence by learning the use of one or more of its words. This obviously works for ‘No bachelor is married’ and the like, and it also works for the basic laws of logic. Anyone who goes counter to modus ponens, or who affirms a conjunction and denies one of its components, is simply flouting what he learned in learning to use ‘if’ and ‘and.’ (I limit this

---

190 Quine (1973), p. 80.
to native speakers, because a foreigner can have learned our words indirectly by translation.)

Quine thus transforms his earlier community-wide account of analyticity into a speaker-relative one: A sentence $S$ is analytic-for-Jones just in case Jones learned to count $S$ true in the process of learning one or more of $S$’s component expressions. This modification yields a behavioral characterization of analyticity in terms of an individual speaker’s language-learning-based dispositions to assent to a sentence.

Quine’s shift from a community-wide characterization of analyticity to a speaker-relative characterization is significant, for it is only if we conceive of analyticity in the former way that we will be inclined to say that no distinction can be drawn between changes of meaning and changes of belief. In conceiving of analyticity in terms of community-wide assent, we abstract away from the workaday interpretive standards used by individual interlocutors to navigate particular discursive interactions. And when we transcend concrete conversations in this way and attempt to consider sentences from an a-contextual perspective, it no longer makes sense to invoke the distinction between meaning change and belief change. For the very point of drawing such a distinction is to facilitate interpretive coordination in particular, concrete contexts of discursive interaction. Quine, as we have seem, needs a way to preserve the distinction between meaning-change and belief-change, so it is no wonder he reformulates his earlier account in speaker-relative terms. From here on out, I will focus on Quine’s speaker-relative notion, which I’ll call Q-analyticity.

In characterizing Q-analyticity as he does (i.e., in terms of speaker-relative language-learning-based assent-dispositions), Quine is deliberately abandoning what in earlier chapters I called analyticity’s metaphysical role (i.e., explaining necessary truth) as well as its epistemological

---

192 Ibid.
role (i.e., grounding a priori knowledge). Instead, Q-analyticity is designed to accommodate analyticity’s interpretive role—that is, Q-analyticity promises to give us a way to make sense of the distinction between change of meaning and change of belief and thus the distinction between verbal and non-verbal disputes.

According to Quine, “truths deducible from analytic ones by analytic steps would count as analytic in turn,” and therefore “the logic of truth-functions, quantification, and identity” would all count as Q-analytic on his view. This, in effect, is Quine’s account of the apparent obviousness of basic logic and mathematics: To say that a basic arithmetical or logical truth is obvious (for Jones) is to say that Jones learned the use of its component terms by learning to assent to it, or else derived it from truths that were learned in this way.

Despite consistently taking the law of non-contradiction to be Q-analytic, Quine seems to change his mind about the status of the law of excluded middle. In *Roots of Reference*, he claims that “excluded middle, which [intuitionists] do contest, is not similarly bound up with the very learning of ‘or’ or ‘not’; it lies rather in the blind quarter of alternation.” Later, however, in “Two Dogmas in Retrospect,” he claims just the opposite: “Repudiation of the law of excluded middle would be a change of meaning, and no less a change of theory for that.” Presumably, he admits this because a classical logician (like himself) could derive excluded middle from obvious (i.e., Q-analytic) premises.

As obscure and unclear as Quine says traditional analyticity is, one thing he seems quite sure of is that ‘$S$ is analytic’ always entails ‘$S$ is true’. For example, in “Two Dogmas,” he writes: “Actually we do know enough about the intended significance of ‘analytic’ to

---

193 Ibid.
know that analytic statements are supposed to be true.

Q-analyticity, like stimulus-analyticity, is defined in terms of assent—the behavioral manifestation of belief or holding true. If sentence $S$ is Q-analytic for Jones, then it follows that Jones must believe that $S$ is true (i.e. be disposed to assent to it) and must have acquired that belief in the process of language-learning. So, if at some point Jones loses his disposition to assent to $S$, then $S$ no longer counts as Q-analytic for Jones. But this result is problematic. For suppose that as a child Jones is taught to count ‘All Germans are boche’ and ‘All boche are prone to cruelty’ as true in the process of learning the use of ‘boche’. Now suppose that later on, after Jones realizes that the empirical evidence fails to support these generalizations, he loses the relevant assent-dispositions. Even without these dispositions, Jones will still understand the expression ‘boche’ perfectly well and will for that reason consider the move from ‘X is boche’ to ‘X is prone to cruelty’ an obvious step.

Moreover, since only beliefs acquired in the process of learning to use words count as Q-analytic, it is hard to see how Quine can account for the fact that a certain principle might come to seem obvious to Jones, despite his coming to believe it long after he first learned to use its component expressions. For example, suppose Jones is taught how to use the word ‘table’ as a child by being presented with various tables and learning to count ‘Tables are furniture’ as true outright. Years later, after reading too much analytic metaphysics, Jones becomes convinced of the self-evident obviousness of the sentence ‘If there are simples arranged table-wise in region $R$, then there is a table in region $R’$. He has come to find it obvious; but this sentence will not count as Q-analytic for Jones.

---

196 Quine (1951), p. 34.
4.5 Constitutivity Commitments and Linguistic Understanding

Although sufficiently behavioral, Q-analyticity is defined in terms of the wrong sort of behavior. Something more subtle than assent is wanted here—and this is what my account of interpretive coordination provides. On my account, the behavioral dispositions relevant to determining which sentences a person treats as constitutive are her dispositions to react to the denial of certain sentences, modulo the current information state of a particular discourse interaction. The reactions in question involve the deployment of precisely the sort of intensional vocabulary—talk of meaning, synonymy, necessity, essence, etc.—that Quine is characteristically skeptical about.

The terms of a natural language are shared by a community, and competent speakers are expected to use those terms correctly.\textsuperscript{197} Having the ability to use a word correctly is not simply a matter of being disposed to apply it to whatever it is, in fact, true of. Nor is it simply a matter of being disposed to assent to true sentences featuring the word in question. Rather, to be able to use a word correctly is to know how to use it in accordance with its meaning, with full awareness that this is what one is doing. A parrot trained to reliably utter the word ‘apple’ whenever presented with an apple does not understand the English word ‘apple’. To be sure, certain parrots may be reliable indicators of particular environmental conditions, and their behavior may thus be described as being in accordance with various rules—but such creatures have no further disposition to positively or negatively assess behaviors that do or do not accord with the rule. It is for this reason that the parrot cannot be said to follow any rules. Indeed, the crucial difference between competent language-users and well-trained parrots lies in the fact that language-users are rule-followers.

\textsuperscript{197} See Rosenberg (1976).
What the language-user has and the parrot lacks are dispositions to interpret expressions in accordance with particular standards of correct use. Having such dispositions means being prepared to distinguish between, on the one hand, pieces of verbal behavior that count as genuine uses of expression $E$; and on the other hand, $E$-tokenings which fail to accord with $E$'s meaning and thus amount to mere noise. This sorting procedure can be done privately, but we can also publicly indicate our interpretive commitments by employing distinctive vocabulary (e.g., by making claims about meaning or essence). The deployment of such vocabulary (or some functional equivalent) by a competent language-user provides prima facie evidence of the constitutivity commitments she currently holds. If there is no sentence such that she is disposed to deploy this vocabulary in response to its denial, then we have no evidence that she knows how to discriminate between correct and incorrect usage. Since having this capacity is essential to linguistic understanding, we may legitimately withhold attribution of competence to someone who displays no behavior whatsoever of this sort. On the other hand, since normally socialized human beings will develop this capacity as a matter of course, we can safely presume our interlocutors to be minimally competent by default with a vast array of common expressions.

I call our practical dispositions to interpret words in accordance with their meanings constitutivity commitments. Roughly, to undertake a constitutivity commitment with respect to expression $E$ is to be disposed to assess $E$-behavior in accordance with a particular interpretive standard. Having such standards at the ready is important because interpreters need to be able to distinguish genuine usage from nonsense, detect confusion or equivocation, decode malapropisms, and so on. The following passage from Wittgenstein helps illuminate the point:
Let us remember that there are certain criteria in a man’s behaviour for the fact that he does not understand a word: that it means nothing to him, that he can do nothing with it. And criteria for his ‘thinking he understands’, attaching some meaning to the word, but not the right one. And lastly criteria for his understanding the word right. In the second case one might speak of a subjective understanding. And sounds which no one else understands but which I ‘appear to understand’ might be called a “private language”.198

Here, Wittgenstein describes three different sorts of behavioral criteria associated with understanding a linguistic expression: (i) behavior that indicates a total lack of understanding; (ii) behavior that indicates a mistaken understanding; and (iii) behavior that indicates correct understanding. My view is that the constitutivity commitments held by the participants in a conversational context provide the standards for determining what sorts of behaviors should count as satisfying (i), (iii), and (iii).

So how can I tell when my interlocutor is treating a certain sentence as a constitutive principle for one of its component expressions? The evidence comes from observing that person’s verbal behavior in order to get a sense of how she is disposed to apply the expression in question. But the verbal behavior we need to observe cannot be mere assent; it must be a more complex sort of response. For a behavioral criterion characterized in terms of assent is simply too coarse-grained to handle important interpretive phenomena that any adequate account of linguistic understanding must address. In particular, a criterion characterized in terms of assent sheds no light on our grasp of inconsistent concepts (whose constitutive principles are inconsistent), nor can it help account for incomplete competence (where a non-expert lacks some, though not all, of the relevant assent-dispositions on the basis of partial ignorance or misinformation), nor the possibility of competent-yet-deviant theorists (experts who hold non-standard theories of the subject matter and thus lack certain relevant assent-dispositions). These three phenomena demonstrate the need for a more subtle

198 Wittgenstein (1953), §269.
account, since they show that treating a sentence as a constitutive principle comes apart from believing that it is true (i.e., being disposed to assent to it).

We should thus characterize the behavioral criterion for distinguishing constitutive from non-constitutive sentences as follows: If participant $U$ in conversation $C$ treats sentence $S$ as a constitutive principle at time $t$, then $U$ must be disposed at $t$ to deploy characteristic intensional vocabulary in distinctive ways in response to the denial of the sentence in question. If $U$ displays no such behavior when prompted, then we have no evidence that she has undertaken the relevant constitutivity commitment. It is always possible, of course, that one’s interlocutor is suppressing the relevant behavior in any given case. It will not make sense, however, to say that $U$ takes $S$ to be a constitutive principle for $E$ even though she would never exhibit the relevant behavioral reactions, even in the most favorable of conditions.

The practice of sorting constitutive from non-constitutive uses is not merely convenient, it is arguably essential to our very ability to use language in the way that we do. If we did not enforce a distinction between what is part of the meaning of a term and what is not, we would constantly be talking past one another—if we could even be said to be “talking” at all. In such a community, no one would distinguish between correct and incorrect verbal performances, so no one would be under any normative pressure to conform to the usage-patterns of her peers. Regularities of usage might arise by accident, but they would not be enforced or maintained by the community. It is not clear that what we think of as “linguistic instruction” would even be possible among such creatures, since ex hypothesi nobody has any dispositions to encourage correct use or discourage incorrect use. At best, such “instruction” would consist in training individuals to reliably emit a certain
sound under certain environmental conditions—a process not significantly different from calibrating a thermometer. The upshot is that without the practice of enforcing a constitutive/non-constitutive distinction, human communication would be limited to mere signaling.

4.6 Concluding Remarks

The practice of distinguishing between constitutive and non-constitutive uses of language makes interpretive coordination possible among individuals with divergent perspectives, which in turn makes possible the sharing of information among interlocutors in a discourse interaction. Constitutivity is not a causal-explanatory property—that is, we should not expect it to play an essential role in our best scientific theories—but this is far from sufficient grounds for abandoning the notion. At the end of the day, what matters for explanatory purposes are the behavioral dispositions I have called constitutivity commitments. Constitutivity commitments are detectable via observation, and the capacity to detect them is something every competent language-user needs to cultivate.

My account of interpretive coordination in terms of the pragmatic negotiation of shared constitutivity commitments is intended to provide an explanation of what we are doing when we deploy the family of intensional concepts Quine raises doubts about in “Two Dogmas.” The strategy is to explain the existence and persistence of the target discourse by describing the underlying pragmatic function it serves, i.e., the job it does in the lives of the creatures who engage in it. It is important to note, however, that explanation is not justification: A story about why we tend to engage in a certain sort of discourse is not automatically a reason in favor of our continued engagement in it. We might, after all, come
to realize that the discourse in question serves a purpose that is in some way illegitimate or unacceptable. But although such explanation does not immediately justify the practice, it can in many cases reveal underlying functions that clearly are beneficial for creatures like us, given what we know about human capacities, needs, and limitations.

Insofar as engaging in moral discourse makes it possible for groups of human beings with conflicting interests to live more harmoniously together, it makes good sense for humans to continue using moral vocabulary. A similar point can be made about our practice of sorting constitutive from non-constitutive uses of language. The vocabulary with which we make our constitutive commitments explicit earns its keep precisely because human discursive practice would be immensely impoverished, if not impossible, without it. And on the assumption that rational conversation and communal inquiry are things to be valued, we have good reason to carry on distinguishing between sentences that are to count as constitutive and those that are not.
Chapter 5

Expressivism, Deflationism, and Constitutivity

5.1 Introduction

The discursive utility of intensional discourse consists in providing human beings a means by which to expose their interpretive dispositions to rational appraisal, thereby facilitating the alignment of those commitments across the speech community. Much like moral discourse, discourse about meaning takes assertoric form, which is why both count as instances of rational discourse more generally. In this chapter, I argue that the best way to accommodate such data is to endorse a thoroughly deflationary version of truth-conditional semantics and formulate a positive expressivist thesis in non-semantic terms. I then try to make the case that the apparent consequences of such a maneuver—i.e., the dissolution of what Robert Kraut calls the Bifurcation Thesis—are not as worrisome as one might be inclined to think. In particular, I suggest a way in which the advocate of global pragmatism can preserve the important explanatory contrast at the heart of the Bifurcation Thesis, without lapsing back into representationalism.
5.2 Reversing the Customary Order of Explanation

There appears to be a fundamental difference between, on the one hand, a sentence like ‘A brother is a male sibling’ and, on the other hand, a sentence like ‘There are tigers in India’. But what best explains this intuition? A traditional strategy invokes the analytic-synthetic distinction: The former sentence is distinctive because it expresses a truth of meaning—necessary and \textit{a priori} knowable—while the latter expresses a contingent truth about the empirical world. On this sort of story, the former is necessary and knowable \textit{a priori} because ‘brother’ has the same meaning as ‘male sibling’, and knowing this is not a matter of collecting empirical evidence; it is simply a matter of grasping the relevant meanings. Intuitive considerations about synonymy, definition, and basic linguistic competence thus seem to suggest that sentences like ‘A brother is a male sibling’ and ‘There are tigers in India’ appear fundamentally different because they are fundamentally different—that is, they express propositions with fundamentally different modal and epistemic profiles, and these differences are grounded in and explained by facts about meaning.

It is important to note that this sort of explanatory strategy requires certain basic ingredients. The requisite inventory includes not only human beings, their linguistic behaviors, and the physical environments in which they are embedded—it also includes an essential explanatory appeal to the \textit{meanings} of shared linguistic expressions. On this view, creatures like us are capable of knowing (and linguistically communicating) various truths, analytic and synthetic, because (i) we are acquainted with objects and events in our physical environment, (ii) we are acquainted with the meanings of the linguistic expressions we competently employ, and (iii) those meanings fix the proper application conditions of the relevant expressions in all possible circumstances.
According to some inflationary views, meaning-properties are non-natural abstract entities which, although otherwise causally inert, are epistemically accessible to us via a special faculty of rational intuition. There are also inflationary views which attempt to ground facts about meaning in more terrestrial matters (e.g., communicative intentions, causal-historical chains of usage, stipulative episodes or baptismal ceremonies, linguistic deference patterns among experts and non-experts, etc.). In either case, ‘meaning’ is assumed to be genuine theoretical term, such that meaning facts are just those facts—whatever they may turn out to be—that serve as the truth-makers for (the weighted majority of) our folk-platitudes about meaning. As such, the core commitment that unifies the inflationary camp is the belief that meanings have a fundamental explanatory role to play in a philosophical theory of language. On such views, competent speakers’ actual and possible usage is explained by their grasp of meanings, and each word’s meaning fully determines how it is correctly applied.

But this way of looking at things is by no means obligatory. One may instead embrace a reversal of the order of explanation implicitly assumed by such strategies. According to this alternative sort of approach, meaning is best explained in terms of human linguistic practice, not the other way around. The principle motivation for reversing the customary order of explanation in this way is the suspicion that the explanatory role traditionally assigned to meaning simply cannot be realized. For if meanings are to play this role, our grasp of them must both (i) explain why competent speakers actually use words as they do, and (ii) provide an objective standard of correctness that determines how those words should be applied in any possible circumstance. But despite its intuitive fit with the phenomenology of competent language use, this classical picture is deeply problematic. The
trouble is that our history of past usage and our linguistic intentions are entirely compatible with many different ways of “going on” in the future. This means that no facts about us—about our behavioral dispositions, our mental states, or anything else— uniquely determine how our words should be applied in novel circumstances.¹⁹⁹

One way to react to such considerations is to insist that there must be some fact or set of facts which can play the required role. Another way to react is to embrace some kind of radical skepticism according to which none of our words means anything. A third way to react is to re-think linguistic correctness itself, rejecting the classical picture in favor of a more plausible alternative. To reject the classical picture is to hold that even though the conditions for the correct application of our terms are unfixed or unsettled, it is still possible to use words correctly (or incorrectly). That is, whether a particular tokening of expression $E$ counts as a correct use of $E$ does not depend on whether there are facts that uniquely determine how $E$ should be applied in all possible circumstances.

Once we reject classical realism about linguistic normativity, space opens up for a pragmatist alternative. The practice of making assessments of linguistic correctness is useful for creatures like us not because it puts us into contact with determinate meanings, but because it makes interpretive coordination possible among individuals with divergent perspectives. This practice is important because in the absence of sufficient interpretive alignment among interlocutors, beliefs cannot be shared and knowledge cannot be transmitted. Indeed, without interpretive coordination, there can be no communal inquiry—no way of reaching a consensus about the way the word is.

¹⁹⁹ See Wittgenstein (1953), §§138-242. See also Kripke (1982) and Wilson (1994). For a (methodologically) behaviorist rout to essentially the same conclusion see Quine (1960), especially chapter 2.
5.3 Expressivism and the Frege-Geach Lesson

The speech act of assertion is a crucial move-type in the discursive game because it allows interlocutors to make public their doxastic commitments and thus subject them to rational scrutiny. One is criticizable if one asserts things that one does not actually believe, or if one lacks sufficient evidential support for the assertions one makes. But competent assertoric practice aims not merely at the expression of sincere, justified, opinion—it aims, in addition, at objectivity or truth. By asserting a particular indicative, a speaker endorses its propositional content, indicating publicly that what it says is correct or accurate by the speaker’s lights. Yet truth outstrips warranted assertibility; for even when our assertions are both sincere and well-justified, we may still be in error if what we have said fails to be true. Each competent speaker is thus entitled to assert ‘Sentence $S$ is true’ whenever she is entitled to assert sentence $S$ itself (and vice versa). The schematic disquotational principles that govern competent use of the truth-predicate—\( \text{From } 'S', \text{ infer } "'S' \text{ is true}" \) and \( \text{From } "'S' \text{ is true}", \text{ infer } 'S' \)—are available as free moves to any genuine participant in the assertion game.

Following Frege, Peter Geach famously points out that the sort of indicatives typically cited as requiring expressivist treatment (e.g., ethical sentences) clearly admit of negation, conditional embedding, and participate in valid inferences. He notes that ethical indicatives can also stand in logical relations to non-ethical indicatives, and may appear in unasserted contexts.\(^{200}\) At bottom, the Frege-Geach point turns on the simple observation that ethical sentences behave, logically and syntactically, just like non-ethical sentences. For traditional Fregeans, this data is easily accommodated. For the non-cognitivist, however, it is a source of seemingly endless difficulty and strife.

\(^{200}\) See Geach (1960), (1965).
Suppose that an ethical sentence like ‘Lying is wrong’ derives its content from a non-cognitive mental state (e.g., the speaker’s disapproval of lying). What happens when this sentence is embedded in an unasserted context? To accept a conditional with an ethical antecedent is not to manifest approval or disapproval of anything. So what, exactly, is the content of a sentence like ‘If lying is wrong, then it’s wrong to get your little brother to lie’? To answer such challenges, the non-cognitivist must develop a whole new semantics for the non-cognitive attitudes and show how the logic is supposed to work.\(^{201}\) To be adequate, this non-cognitivist semantics will have to be recursive and compositional, and must in general mirror the basic logical and syntactic trappings of standard truth-conditional semantics.\(^{202}\)

Moreover, if, as the non-cognitivist holds, the content of ethical sentences is fundamentally different in kind from the content of non-ethical sentences, then there is a further puzzle about how those different kinds of contents—one truth-conditional and the other not—could interact in the required ways. For example, suppose I see Johnny kick his dog and in response I assert ‘Johnny just did something wrong’. When you ask me to defend my claim, suppose I assert ‘Johnny’s action was wrong because it caused his dog pain’. I thus offer you what looks like a valid argument in support of my first assertion. But note that one of my premises is a non-ethical sentence—‘Johnny’s action caused his dog pain’—which according to the non-cognitivist has a fundamentally different kind of content from ethical sentences. If the argument is valid, then it is truth-preserving: Given all true premises, the conclusion’s truth follows. But how is truth to be preserved when only one of the premises

---

\(^{201}\) Classic treatments occur in Blackburn (1984), (1993), (1998), and Gibbard (1990), (2003). For the most well-worked out (and most complex) version of mentalistic expressivist semantics, see Schroeder (2008).

\(^{202}\) Various problems have arisen for Blackburn and Gibbard on this score. See van Roojen (1996) for a critique of their handling of conditionals and Unwin (1999) for a critique of their handling of negation. These and related problems are the central focus of Schroeder (2008).
is a factual statement? Indeed, how could any argument establish the truth of a non-factual (because non-cognitive) conclusion?

Expressivism, especially in its earlier historical incarnations, is typically formulated in negative terms: an expressivist about discourse D holds that D-indicatives do not describe the world, or do not state facts, or are not truth-apt, or lack truth-conditions. But the negative thesis is not the only aspect of expressivism; there is also a crucial positive thesis: D-indicatives serve to express attitudes, manifest stances, give voice to behavioral dispositions, or the like. Standard versions of expressivism are local views directed at some discourses (e.g., moral discourse) but not others (e.g., scientific discourse).

The Frege-Geach observations are clearly incompatible with the expressivist’s negative thesis, but they are not obviously in conflict with the positive thesis. This suggests that if the expressivist could find a way to keep the positive thesis while dropping the negative thesis, then she could simply side-step this perennial thicket of problems altogether and help herself to the standard Fregean semantic apparatus. In order to be entitled to that apparatus, yet retain her commitment to the pragmatist order of explanation, she needs a way of domesticating the (apparently) metaphysically-loaded vocabulary of truth-conditional semantics. And indeed, there is a way to do this: the expressivist ought to embrace a thoroughgoing deflationism about the family of notions essential to truth-conditional semantics, including truth, contents (i.e., propositions), facts (i.e., true propositions) and properties.²⁰⁴

---

²⁰³ This way of carving things up comes from Price (2011), p. 88. See also Kraut (1990), pp. 159.
²⁰⁴ Price (1994), (2004), recommends this move for similar reasons. See also Rorty (1986) and Williams (1999).
5.4 Deflationism and Explanatory Role

There are many varieties of deflationism about truth-talk, but the core commitment uniting all such views is the idea that the notion of truth plays no explanatory role in a philosophical theory of language—its discursive utility wholly consists in serving a non-explanatory function (e.g., as an expressive device for opaque endorsement and generalization). A language that contains a truth-predicate has vastly increased expressive resources in comparison with a language without one. Equipped with a truth-predicate, one can endorse infinite conjunctions (e.g., ‘Every sentence of the form “P and not-P” is false’), and one can agree with sentences one does not understand (e.g., “Schee ist weiss” is true’ uttered by a monolingual English speaker).

Deflationists hold that you know all there is to know about truth once you know how to infer in accordance with the basic disquotational principles that govern the truth-predicate. There is no further philosophical question about the nature of truth. The deflationist’s opponent—call her the inflationist—is someone who takes truth to play a crucial explanatory role in an adequate account of human linguistic practice. An inflationist will be tempted to ask what sorts of facts or states-of-affairs play the role of truth-makers for our claims. She may pursue such questions via conceptual analysis, producing content-specifying biconditionals that purport to explain why various sentences are true. The deflationist, by contrast, sees such biconditionals as theoretically uninteresting.

---

205 Some important examples of the tradition include Ramsey (1927), Ayer (1935), Quine (1986), Horwich (1990) and Brandom (1994).
206 Armour-Garb (2012) suggests that the truth-predicate is able to serve as a device for opaque endorsement and generalization because it is, in a broader sense, a device for semantic decent—its job is to “undo some nominalization of a content-vehicle” (p. 275).
Any deflationary approach to truth needs two aspects: (a) an account of the \textit{function} of the truth-predicate; and (b) an account of the \textit{mechanism} by which that function is performed.\footnote{Williams (1999), pp. 547-548.} To be a deflationist (in the broad sense) is to hold that the function of truth-talk is entirely non-explanatory. Where the various deflationary approaches differ is in their account of the mechanism by which the function is performed: Quine says ‘is true’ is a device for disquotation (i.e., for cancelling quotation marks), Paul Horwich transposes the Quinean mechanism and treats ‘is true’ as a predicate of propositions, and Robert Brandom takes ‘is true’ to be an operator for forming prosentences which inherit their content anaphorically.\footnote{See Quine (1986), Horwich (1990), and Brandom (1994).} For present purposes, the particular details of how a deflationist should describe the workings of the mechanism are not immediately relevant. My focus will be on the methodological stance that unifies deflationists—i.e., their rejection of explanatory appeals to truth.

The further notions employed in truth-conditional semantics—\textit{proposition}, \textit{property}, and \textit{fact}—are intimately linked to \textit{truth}, forming an intensional circle of and interdependent and interdefinable notions. Given these constitutive connections, one should either take a deflationary approach to the entire family, or reject deflationism for the whole family. That is, if one takes a deflationary approach to truth-talk, one should also take a deflationary approach to fact-talk, property-talk, and proposition-talk. A disquotationalist like Quine might want to avoid this move, since he takes truth to be a property of sentences and would in any case be uncomfortable with talk of propositions. Such discomfort is misplaced, however. We should conceive of the essential vocabulary of the apparatus of truth-conditional semantics as implicitly defined by sets of transformation rules, much like the
sentential connectives are implicitly defined by their introduction and elimination rules. On such a conception, talk of properties and propositions becomes no more ontologically worrisome than talk of conjunction or negation.

The thoroughgoing deflationist about this whole family of notions will agree that the German sentence ‘Es regnet’ is correctly translated as ‘It’s raining’ in English, and she will also agree that from this it follows that the two sentences express the same proposition. She will further agree that ‘Fido and Lassie are dogs’ entails ‘Fido and Lassie share a property, namely the property of being a dog’. However, she will deny that claims about correct translation are to be explained in terms of relations between abstract Fregean senses, and she will deny that claims about properties are to be explained in terms of relations between concrete particulars and abstract universals. Such inflationary suggestions are, from this point of view, nothing more than self-congratulatory pats on the back which can neither explain nor justify linguistic practice.

Note that once we go deflationist, we are no longer entitled to claim that any of the D-indicatives we are willing to assert are non-descriptive, non-fact-stating, non-truth-apt, or lacking in truth-conditions. For to assert an indicative just is to commit oneself to its truth, so any attempt to deny its truth or truth-aptitude amounts to a retraction of the original assertion. Moreover, what goes for truth goes for facts as well: If I’m authorized to assert a certain indicative, I’m equally authorized to call it true or to say that what it states is a fact (and vice versa). This is why it makes no sense to assert ‘Grass is green’, yet deny ‘It is a fact that grass is green’—if you are willing to assert the former, you ought to be willing to assert the latter.
Deflationism does not, however, undermine the expressivist’s positive thesis regarding the special function served by sincerely asserted D-indicatives. The result of combining deflationism with the positive thesis is thus a version of expressivism without non-factualism—a view compatible with (deflationary) Fregean semantics, but also compatible with a rejection of Fregean (i.e., classical realist) metaphysics of meaning. This move makes available a thoroughly pragmatist theoretical perspective in which semantic notions are denied their traditional explanatory role, but in light of their non-explanatory utility, are not eliminated altogether. In place of those semantic notions, we appeal instead to a pragmatic or use-theoretic vocabulary as our basic theoretical vocabulary (i.e., the sort of vocabulary used to describe, vis-à-vis the positive thesis, what we are doing when we assert D-indicatives). By recasting the expressivist’s positive proposal in non-semantic terms and adopting a wholly deflationary truth-conditional semantics, a new kind of pragmatist approach to linguistic theorizing becomes available.\(^{209}\)

The sort of expressivism that emerges when we reject the negative thesis, retain the positive thesis, and fully embrace deflationism, is a necessarily global, rather than a merely local, view.\(^{210}\) Against this backdrop, the theoretical task becomes that of distinguishing regions of discursive practice on the basis of their different pragmatic functional roles rather than their contents. On this picture, moral discourse differs from scientific discourse not because they have fundamentally different sorts of content, but because the alignment of moral attitudes is useful to human beings in a different (though no less legitimate) way than the alignment of scientific attitudes is. Communal alignment of scientific commitments is

---

\(^{209}\) See Price (2004) pp. 214-217. As Price puts it, “A modest Fregean content-specifying theory is compatible with the view that the interesting theoretical vocabulary for linguistic theory is pragmatic or use-theoretic, rather than semantic.” (p. 215)

\(^{210}\) See Macarthur and Price (2007).
useful because it facilitates reliable prediction and causal explanation of natural phenomena. By contrast, communal alignment of moral commitments is useful for a distinctively non-explanatory purpose: it facilitates social harmony and maintains the integrity of the social contract. The global pragmatist’s charge is to pursue, throughout assertoric discourse, similar anthropological or genealogical accounts couched in use-theoretic vocabulary.

The adoption of globalized pragmatism not only provides a way to side-step the Frege-Geach problem, it also provides a way to avoid the familiar objection that expressivism about meaning-talk is an incoherent or internally unstable position. The objection turns on the allegation that the expressivist about meaning-talk is automatically committed to an incoherent form of semantic non-factualism. However, since the global pragmatist takes descriptiveness to be an essential property of every assertable indicative, she does not count as a non-factualist about any assertoric region of discourse. Thus the instability objection fails. This result removes important conceptual barriers that otherwise might appear to preclude a pragmatist account of interpretive coordination from getting off the ground.

5.5 Constitutivity Commitments in Interpretive Practice

Consider the sort of vocabulary typically mobilized when we conduct linguistic instruction, pursue interpretive coordination, or assess linguistic understanding. Paradigmatic examples include: (a) broadly meta-linguistic claims about word meaning or definition, analyticity attributions, claims about conceptual connections or semantical rules, etc.; and (b)

---

211 See Boghossian (1990).
broadly object-level claims about the nature or essence of a thing, attributions of
metaphysical necessity, claims about what it is to be a certain object or kind, etc.

Although (a)-type claims have different content from (b)-type claims—the former
are about language while the latter are not—both sorts of claims play the same crucial
pragmatic role. In particular, (a)-type and (b)-type claims function to express the
constitutivity commitments undertaken by various interlocutors in a context. Constitutivity
commitments guide interpretation by providing conversational participants with *prima facie*
standards by which to distinguish changes of meaning from changes of belief. If your
interlocutor says something that implies the negation of a sentence you take to be a
constitutive principle, this is *prima facie* evidence that the speaker does not mean what you do
by some of the words she is using.

Competent speakers use these constitutive principles as interpretive standards for
deciding whether a certain piece of verbal behavior counts as (i) a correct use of language, (ii)
an incorrect use of language, or (iii) nonsense. On this view, to be a competent language-user
is to have the ability to sort (apparently linguistic) behavior in this way and to be disposed to
courage behavior one takes to be correct and discourage behavior one takes to be
incorrect. What counts as competent usage is not determined in advance by the intrinsic
nature of non-human reality, but is instead a matter of negotiation between rational
interlocutors who wish to pursue communal inquiry in accordance with shared goals. This
account of linguistic understanding thus stands in stark contrast to the traditional classical
realist picture according to which to be a competent language-user is to be *en rapport* with
practice-transcendent non-human entities (e.g., reified Fregean senses, or Plato’s Forms, or
Ideas in the Mind of God, or whatever).
In addition to their (deflationary) truth-conditional content, sincere constitutivity claims of both the (a)-type and the (b)-type also convey pragmatic information about the speaker’s interpretive dispositions. Someone who sincerely asserts ‘Being a feline is an essential property of being a cat’ says that being a feline is an essential property of being a cat. But in producing this utterance she also indicates that she has undertaken the corresponding constitutivity commitment and thereby licenses certain interpretive expectations on the part of her audience. Note, however, that since saying is not the same as signaling (indicating), constitutivity claims are not mere self-reports of the speaker’s interpretive dispositions. Moreover, the pragmatic information conveyed by sincere constitutivity claims is not part of the truth-conditional content of those claims, nor should it be thought of as a second kind (or dimension) of content.

As a general rule, it is appropriate to make a constitutivity claim only if one actually has the relevant interpretive dispositions (i.e., those that typically cause competent speakers to make the constitutivity claim in question). This helps explain why human language-users have access to claims with such content in the first place: the point of having such contents available is to allow us to expose our constitutivity commitments to rational scrutiny, thus facilitating interpretive coordination and promoting communal regularities of usage. In this way, we get a deeper understanding of the pragmatic factors that underlie disputes over a word’s meaning or a thing’s essence.

For example, consider the sentence “The word “cat” means “domestic feline”” and its negation ‘It is not the case that “cat” means “domestic feline”’. Someone who asserted both of these sentences would be making a mistake. But what explains the source of that mistake? It is not enough to simply note that the speaker has said something of the form ‘p and not-p’
and then proceed to rehearse the introduction and elimination rules for classical ‘~’. Nor is it enough to locate the mistake in the fact that any claim of the form ‘p and not-p’ expresses a contradictory proposition, since to say that a claim expresses a contradictory proposition just is to say that it has (or is equivalent to a sentence with) the form ‘p and not-p’. These diagnoses are not wrong or false—but nor are they properly explanatory. The pragmatist orientation makes it possible to say something more interesting about the source of such mistakes. It is ultimately in virtue of a fundamental *practical incompatibility* between interlocutors’ interpretive dispositions that they treat such pairs as logically inconsistent.

### 5.6 The Bifurcation Thesis and Promiscuous Truth

Global pragmatism is incompatible with an important thesis that many will feel is essential to the Humean expressivist tradition. In particular, global pragmatism is incompatible with the *Bifurcation Thesis*: i.e., the thesis that assertable indicatives may be divided into two classes—the descriptive and the non-descriptive—and that the former, but not the latter, have propositional content.\(^{212}\) According to the global pragmatist, to have propositional content just is to be an assertable indicative, which means that descriptiveness turns out to be a grammatical, rather than a metaphysical, feature of sentences.\(^{213}\) According to the merely *local* pragmatist, however, descriptiveness must be understood non-grammatically in order to sustain the Bifurcation Thesis. Thus, we appear to be at an impasse. In this section, I shall try to zero-in on the core concerns that motivate the dispute between the bifurcationist and the global pragmatist and consider whether an ecumenical compromise between the two camps can be arrived upon.

---

\(^{212}\) Kraut (1990), p. 159.  
\(^{213}\) See Price (1994), pp. 67-68.
Since the core notions utilized in truth-conditional semantics are constitutively interconnected with one another, their inferential roles can be made explicit by raising and answering a series of Carnapian “internal questions.” The answers to these internal questions are trivially entailed by the constitutive principles that govern the cluster of notions in the truth-conditional semantics family. Here are some of those questions and answers:

- What is it for a sentence to be descriptive rather than non-descriptive? Answer: A descriptive sentence is in the business of stating facts, whereas a non-descriptive sentence is in some other line of work.

- What is it for a sentence to be fact-stating? Answer: A fact-stating sentence is one that purports to express a true proposition.

- What is it for a sentence to express a proposition? Answer: To say that a sentence expresses a proposition is to say that it possesses truth-conditions.

- What is it for a sentence to possess truth-conditions? Answer: To say that a sentence possesses truth-conditions is to say that it can be expected to interact logically and syntactically with other sentences in certain systematic ways—it can be, for example, negated, embedded in conditionals, and participate in valid inferences.

- What is it for a sentence to express a true proposition? Answer: To say a sentence expresses a true proposition is to say that (i) it possesses truth-conditions; and (ii) it correctly describes the way the world is.

- What is it for a sentence to correctly describe the way the world is? Answer: To say that a sentence correctly describes the world is to say that it correctly ascribes certain properties to certain objects.

- What is it for a sentence to correctly ascribe properties to an object? Answer: To say that a sentence correctly ascribes property $P$ to object $O$ is to say that $O$ is (or has) $P$. 

137
Given this deflationary framework, it follows that anyone entitled to assert ‘Murder is wrong’ is also thereby entitled to assert ‘There is something, namely the property of wrongness, that murder has’. Properties, therefore, come cheap. So whether or not an assertable indicative counts as descriptive apparently cannot be cashed out in terms of whether or not its predicates express properties.

Such results are likely to appear deeply problematic to the bifurcationist. For once we embrace the far-reaching consequences of deflationism, how are we supposed to understand the alleged contrast between the expressivist and her realist opponents? The globalized use-theoretic pragmatist must reject the negative thesis for any assertoric discourse, but one might worry that without some such thesis, the distinction between realists and expressivists will simply collapse. This is the concern that James Dreier calls the problem of Creeping Minimalism:

There is a property for each intelligible predicate, but no ontological commitment involved in accepting these minimal, deflated properties. Once we understand that ‘wrong’ and the like behave logically and grammatically as predicates, there is no further question of whether there is a property of being wrong.

Minimalism sucks the substance out of heavy-duty metaphysical concepts. If successful, it can help Expressivism recapture the ordinary realist language of ethics. But in so doing it threatens to make irrealism indistinguishable from realism. That is the problem of Creeping Minimalism.  

A natural way to address the worry is to find some means of blocking the inference from ‘Sentence S is an assertable indicative’ to ‘Sentence S is descriptive’, thereby preserving the Bifurcation Thesis in its original form. An alternative way to address the worry is to resist endorsing the original thesis and attempt to accommodate the bifurcationist intuition by other means. The former strategy has been pursued by Robert Kraut, and the latter strategy by Huw Price. It will be instructive to compare them.

---

Both Kraut and Price hold that there is space within a deflationist account for an even richer story about the function of the truth-predicate along the lines of the expressivist’s positive thesis. That is, in addition to being a grammatical device for opaque endorsement and generalization, sincere applications of the truth-predicate also serve to manifest a stance or evince a behavioral commitment. Kraut and Price pursue very different projective accounts, but both views qualify as deflationary because they hold that the discursive utility of truth-talk is entirely non-explanatory.

For Kraut, it is precisely the projectivist component of truth-talk that is crucial to the preservation of the original Bifurcation Thesis. He suggests that claims like ‘S is true’ or ‘S has truth-conditions’ serve to express a kind of praise for select indicatives that the speaker deems “worthy of confidence vis-à-vis projects that matter to him at the time.” Since not every assertable indicative is eligible for such praise, the truth-predicate is nonpromiscuous and thus the original Bifurcation Thesis can be maintained. According to Kraut,

>a semantic irrealist needs to tie truth-condition talk to a more robust compliment—a more selective kind of honor—than one automatically deserved by any assertable indicative. To be robust, a conception of truth need not be realist; it need only be nonpromiscuous, allowing that certain indicatives, though attractively assertable, are unworthy of that special status signaled with “x has truth conditions.”

For Kraut, then, praise-eligible indicatives are descriptive, praise-ineligible indicatives are non-descriptive. And what is it to be praise-eligible? Kraut’s answer: A user of the truth-predicate singles out for praise “indicatives that [the] speaker regards as somehow ineliminable in light of his special explanatory concerns.”

---

216 Ibid., p. 259.
217 Ibid., p. 252.
Just recently, Kraut has supplemented and clarified his view by focusing his attention on properties. He proposes a “robust deflationism about properties” in the form of a three-stage recipe. First, at stage one, we begin by noting that some predicates are projectible and others not; that some predicates figure into natural laws and others do not; that some sets exhibit nomological unity whereas others are simply gerrymandered collections. “Realist” theories of properties are customarily invoked to explain such phenomena: projectible predicates are those that express real properties; laws of nature (as opposed to mere universal generalizations) are about properties and magnitudes; members of “natural kinds” share properties. And so on. Properties, when invoked in this setting, purport to play an explanatory role.218

At stage two, we realize that there is reason to doubt that such appeals to properties do genuine explanatory work; rather the appeals to properties merely redescribe in lofty metaphysical terms—with little advance in understanding—the phenomena in question (projectibility, nomologicality, underlying unity among predicate extensions).219

Finally, at stage three, we notice that properties, much like shadows, are dependent entities, the existence and behavior of which depend upon the semantics of predicates, but nonetheless real. This is the core insight of RDP, which insists that some—but not all—predicates cast such metaphysical shadows: those predicates, for example, that play essential roles in explanatory projects. … Only predicates meeting the criterion of explanatory indispensability are said to express properties. RDP thus avoids the result that every meaningful predicate expresses a property, while also avoiding the view that properties provide explanations and/or justifications of linguistic practice. This blocks the spreading infection of creeping minimalism.220

At the end of the journey, we have RDP: sincere assertions of indicatives of the form ‘Predicate “F” expresses property P’ serve to signal the speaker’s commitment to the explanatory indispensability of F-talk.221 Thus, if one is an expressivist about discourse D, then one is committed to holding that assertable D-indicatives feature predicates that fail to express properties, and therefore fail to have truth-conditions (that is, fail to express propositions).

219 Ibid., p. 27.
220 Ibid.
221 See also Kraut (2010).
It should be noted that RDP closely resembles what Kraut says elsewhere about existence-talk. On Kraut’s account of ontological discourse, sincere assertions of indicatives of the form ‘Fs exist’ or ‘There are Fs’ express the speaker’s commitment to the explanatory indispensability of F-talk. Unfortunately for Kraut, the combination of RDP with this account of existence-talk appears to yield some problematic consequences. To see this, consider the following chain of entailments: An assertion of ‘Grass is green’ entails ‘Grass has the property of being green’ which entails ‘There is a property, namely being green, that grass has’, which in turn entails ‘Properties exist’. It thus appears that, on Kraut’s view, a sincere assertion of ‘Grass is green’ commits the speaker, via these entailments, to the explanatory indispensability of property-talk. This result is problematic for two reasons. First, Kraut does not want to say that the predicate ‘is a property’ plays an essential role in explanatory projects; rather, he wants to say that property-talk is dispensable for explanatory purposes. Second, there is a puzzle about how a descriptive indicative like ‘Grass is green’ could possibly stand in an entailment relation to a non-descriptive indicative like ‘Properties exist’ at all.

It might appear that Kraut could avoid the trouble by denying one of the entailments in the chain. But it is not clear that this is an option, since he holds that properties are “dependent entities, the existence and behavior of which depend upon the semantics of predicates.” This characterization suggests that the existence of properties does follow trivially from any true claim in which expressions of the appropriate grammatical type (i.e.,

---

222 See Kraut (forthcoming-a), p. 42.  
223 Kraut is aware of this, noting that “Entailment relations require, at the very least, truth-evaluability of antecedents and consequents; insofar as ontological discourse functions expressively, claims about existence are without truth value and thus cannot enter into logical entailment relations” (ibid., p. 37).
predicates) feature. In other words, ‘Properties exist’ appears to be the trivial answer to a Carnapian internal question.

Let us put such issues aside and return to the question of whether Kraut’s views on property-talk can rescue the Bifurcation Thesis from the threat of Creeping Minimalism. Although RDP is claimed to “block the spreading infection of creeping minimalism,” Kraut does not spell out the details explicitly. Presumably, the reasoning would proceed as follows. In order to preserve the Bifurcation Thesis, the deflationist needs a selective or non-promiscuous truth-predicate. Adopting RDP gives us a principled way to sort assertable indicatives: the ones that are eligible for special praise are those that feature explanatorily indispensable predicates. Thus, if we accept RDP, we can construct a notion of truth which is properly applicable to select assertable indicatives (call these descriptive) but not others (call these non-descriptive).

The success of this line of reasoning depends upon the viability of Kraut’s suggestion that it is appropriate to call an assertable indicative true only if one is committed to the explanatory ineliminability of the vocabulary in which that sentence is couched. But as a description of competent discursive practice, the suggestion looks to be false. After all, there are competent moralizers who, having accepted Blackburn-style quasi-realism, take it to be appropriate to call moral indicatives true while also holding that moral vocabulary is explanatorily inert. The problem for Kraut’s suggestion is that while ‘S is couched in explanatorily indispensable vocabulary’ is plausibly non-promiscuous, it is much harder to make the case that ‘S is true’ is equally selective.

Like Kraut’s view, Huw Price’s brand of deflationism about truth-talk includes a crucial projective component. But where Kraut’s projective component is supposed to help
ground bifurcationism, the projective component of Price’s account is in direct tension with the Bifurcation Thesis. Price argues that calling a sentence true serves to indicate the speaker’s commitment to a norm of assertion stronger than subjective justification, and that all rational discourse must be governed by this norm. He makes his case by considering what a discursive practice that lacked this truth-norm would be like and argues that such a practice would be severely impoverished in comparison with our own.

According to Price, the truth-norm is in play by default because we take disagreements to matter by default. Price says that the truth-norm is a norm which speakers immediately assume to be breached by someone with whom they disagree, independently of any diagnosis of the source of the disagreement. Indeed, this is the very essence of the norm of truth, in my view. It gives disagreement its immediate normative character, a character on which dialogue depends, and a character which no lesser norm could provide. He notes that we sometimes “opt out” of this norm, deciding in particular contexts to allow for no-fault disagreements, but these are the exceptions rather than the rule.

Price invites us to imagine a community—the “Mo’ans”—whose discursive practice is not governed by the truth-norm. (The name derives from their practice of making “merely opinionated assertions.”) Among the Mo’ans, an indicative counts as correctly assertable just in case the speaker (a) sincerely believes it and (b) is sufficiently justified in believing it by her own lights. Competent Mo’an interlocutors are disposed to reprimand those who fail to satisfy one or both of these norms. Suppose Flora and Virgil are Mo’ans. If Flora asserts ‘There are no black swans’ and Virgil asserts ‘Black swans exist’, it does not follow that one of them must be incorrect. For as long as Flora sincerely believes that there are no black swans and takes herself to be justified in that belief, it is correct for her to assert ‘There are

---

no black swans’. This is true even if we suppose that Virgil, as a matter of fact, has actually encountered black swans that Flora doesn’t know about. Any additional information about swans is simply irrelevant to the correctness of Flora’s claim. Given the Mo’ans’ discursive norms, Flora feels no unease about the fact that Virgil said something logically inconsistent with what she said. As a Mo’an, what reason does she have to be motivated to pursue further dialogue? After all, Mo’ans have no disposition to criticize claims that satisfy both Mo’an norms, so from Flora’s point of view, her views could not be “improved” by consulting Virgil, no matter how experienced he is with swans.

Note that the Mo’an situation would remain the same even if we allowed them a disquotational truth-predicate. Price puts the point as follows:

A speech community of this imagined kind could make use of a disquotational truth predicate, as a device to facilitate agreement with an expression of opinion made by another speaker. “That’s true” would function much like “Same again” or “Ditto,” used in a bar or restaurant. Just as “Same again” serves to indicate that one has the same preference as a previous speaker, “That’s true” would serve to indicate that one holds the same opinion as the previous speaker. The crucial point is that if the only norms in play are subjective assertibility and personal warranted assertibility, introducing disquotational truth leaves everything as it is. It doesn’t import a third norm.²²⁶

It is in virtue of this fundamental point that Price advocates a version of deflationism about truth-talk which includes an additional projective component.

There is no doubt that it is useful to be able to assert sentences like ‘All the axioms and theorems of Number Theory are true’ and ‘Johnny said something true yesterday, but I can’t remember what it was’. But the ability to form such constructions, for all its utility, is not all there is to the practice of calling sentences true. According to Price, competent use of our truth-predicate requires a commitment to a third norm: “If Not-P, then it is incorrect to

assert that \( P \); if \( \neg P \), there are prima facie grounds for censure of an assertion that \( P \).\(^{227}\)

The idea is that if we did not bind ourselves by this norm, there would be no social pressure to resolve disagreements in attitude between individuals with different subjective perspectives. The fact that we clearly \emph{do} feel pressure to resolve disagreements shows that we take ourselves to be governed by the truth-norm.

Price admits that the Mo’an community may not even be a coherent possibility at all. He notes that it is doubtful whether such a practice deserves to be called genuinely \emph{assertoric} or \emph{discursive} or \emph{rational}. This is, in the end, the very point that Price’s thought experiment is supposed to establish. The example of the Mo’ans is presented precisely to convince those who do not antecedently doubt its possibility that no rational discursive practice could really work that way. If you doubt the intelligibility of the thought experiment, then you’re on Price’s side from the start.\(^{228}\)

The truth-norm, Price tells us, plays both a “passive” and an “active” role in the lives of human language-users. On the passive side, the truth-norm “creates conceptual space for argument” and makes intelligible “the idea that we might improve our commitments by seeking to align them with those of our community.”\(^{229}\) On the active side, the truth-norm “provides an immediate incentive for argument, in that it holds out to the successful arguer the reward consisting in her community’s positive evaluation of her dialectical position.”\(^{230}\) Thus, on Price’s view the pragmatic function of truth-talk—the deep function that goes beyond its disquotational role—is that of facilitating \emph{dialogue} or \emph{argument} among individuals with divergent perspectives, thereby making possible the improvement of individual

\(^{227}\) Ibid., p. 170.
\(^{228}\) See Price and Rorty (2010), pp. 255-258.
\(^{230}\) Ibid., p. 175.
viewpoints and facilitating convergence of opinion across the community. One’s commitment to the truth-norm manifests itself behaviorally as “a disposition to criticize, or at least disapprove of, those with whom one disagrees.”²³¹

5.7 Bifurcationism and Eleatic Commitment

It is important to note that for all the “realist-sounding” things Price is able to say about truth, his view is clearly not an inflationary account. In keeping with the methodology of deflationism, Price makes no essential explanatory use of the notion of truth. The following passage makes the point vivid:

To use a Rylean metaphor, my view is thus that truth supplies factual dialogue with its essential esprit de corps. As the metaphor is meant to suggest, what matters is that speakers think that there is such a norm—that they take themselves to be governed by it—not that their view is somehow confirmed by science or metaphysics.²³²

As long as you know how to play the game, it really doesn’t matter what sort of metaphysical redescription of the practice you go in for. You could be a raving realist or a hard-core fictionalist, but all that really matters is that you are disposed to criticize those with whom you disagree.

On the other hand, although Price’s view is obviously not inflationary, it does not seem right to call it “irrealist” or “anti-realist” either. Indeed, the account has interesting consequences in this regard. Price writes:

In common with other deflationary approaches to truth, the present account not only rejects the idea that there is a substantial metaphysical issue about truth (a substantial issue about the truthmakers of claims about truth, for example). Because it is about truth, it also positively prevents “reinflation.” In other words, it seems to support a general deflationary attitude to issues of realism and anti-realism. If so, deflationism about truth is not only not

²³¹ Ibid., p. 177.
²³² Ibid., p. 165.
to be identified with fictionalism, but tends to undermine the fictional–non-fictional distinction, as applied in the metaphysical realm.\textsuperscript{233}

The upshot, then, is that we must reject the Bifurcation Thesis. But we need not fear the threat of so-called Creeping Minimalism, however, because the proper reaction is simply to walk away from realist–anti-realist debates altogether.

On this score, Price draws an analogy with two importantly different ways of rejecting theism.\textsuperscript{234} On the one hand, one could reject theism by \textit{using} theological language to deny that a deity exists. On the other hand, one could reject theism by \textit{avoiding} the use of theological language altogether—by refusing to engage in god-talk even to make negative existential claims. The latter way of rejecting theism has important advantages, for in abandoning the vocabulary altogether we simultaneously absolve ourselves of any obligation to defend claims like ‘God does not exist’.

Unlike theological discourse, whose social utility is dubious at best, moral discourse is a practice worth keeping. The same goes for meaning-talk and truth-talk. This assessment is compatible with a realization that the utility of these ways of talking consists in serving non-explanatory functions.

Such considerations lead us back to an important insight that underpins Kraut’s attempt to save bifurcationsim. I submit that the distinction between realists and expressivists has essentially to do with what we might call the \textit{Eleatic commitments} of the two camps—that is, the difference between them seems to consist in the fundamental explanatory resources to which each camp is committed. I think this insight can be retained in light of what Price says about truth and rational discourse, in the following way: One

\begin{itemize}
\item \textsuperscript{233} Ibid., p. 182.
\item \textsuperscript{234} Ibid., p. 181.
\end{itemize}
doesn’t count as a realist about discourse D unless one thinks that the practice cannot be adequately explained without essential use of D-vocabulary. By contrast, the expressivist about discourse D thinks that an adequate explanation of the practice requires that we merely mention, rather than use, D-vocabulary. We can thus think of a realist/expressivist dispute over discourse D as a dispute over the proper methodological constraints that ought to govern adequate theorizing about D—that is, a dispute over which Eleatic commitments theorists ought to endorse with respect to D. 235

5.8 Concluding Remarks

Globalized pragmatism is the result of dropping the negative thesis from Humean expressivism and combining the positive thesis with Price’s version of deflationism about truth. The theoretical framework thus provided naturally makes room for an expressivist account of interpretive coordination in terms of the negotiation of constitutivity commitments.

This combination of views is also compatible with the apparatus of truth-conditional semantics—that is, the formal machinery and concepts developed in the Fregean tradition. However, the way this apparatus is implemented also matters, since certain ways of implementing it might require inflationary truth. The good news is that there are theories available which do not require inflationary truth and would therefore be compatible with both global pragmatism and with my account of interpretive coordination. One such theory is Donald Davidson’s interpretivism, 236 and another such theory is Robert Brandom’s

235 For further discussion of the constraints on vocabulary-specific deflationary and expressivist accounts, see Williams (2011).
236 See the essays in Davidson (1984). See also Williams (1999).
normative inferentialism. These accounts require only a deflationary notion of truth, and both are compatible with the rejection, on Quinean and Wittgensteinian grounds, of classical realism about meaning.

Because embracing globalized pragmatism requires taking the truth-predicate to be promiscuous, orthodox bifurcationism must be rejected. The bifurcationist intuition can be salvaged, however. For the fundamental difference between realism and expressivism turns on whether or not the target vocabulary should be used or merely mentioned when giving a theoretical account of it. Realists are unsatisfied with explanations that merely mention the terms in question, and expressivists are unsatisfied with explanations that use the terms in question. When thus recast in non-metaphysical, non-semantic terms, the realist/expressivist debate is revealed to be rooted in a disagreement over what ought to count as an adequate explanation of certain chunks of human linguistic practice. This move allows us to preserve the important explanatory contrast implicit in the Bifurcation Thesis without feeling compelled to introduce unnecessary complications into our compositional semantics.

---

Bibliography


APPENDIX A: Discourse Structure with Constitutivity Commitments

Scoreboard of a Rational Discourse Interaction $D$:
At any given point $t$ in $D$, the information shared by the interlocutors is structured as follows:

$I$, the set of interlocutors at $t$

$G$, a set of sets of goals in effect at $t$, such that
for all $i \in I$, there is a (possibly empty) $G_i$ which is the set of goals to which $i$ is committed at $t$ to trying to achieve, and

$G = \{ G_i \mid i \in I \}$.
$G_{com} = \{ g \mid \forall i \in I: g \in G_i \}$, the set of the interlocutors' common goals at $t$.
$G_Q = \{ g \in G_{com} \mid$ there is some $Q \in \text{QUD}$ and $g$ is the goal of answering $Q$\}.

For all $i \in I$, if $i$ is a sincere, competent and cooperative interlocutor in $D$, we can use $G_Q$ to characterize two kinds of publicly evident goals held by $i$ (at time $t$):

**Discourse Goals of $i = G_Q$**
**Domain Goals of $i = G_i \setminus G_Q$**

$G_{com} \setminus G_Q$: the set of common Domain Goals of all the interlocutors

$M$, the set of moves made by interlocutors up to $t$, with distinguished sub-sets:

$A \subseteq M$, the set of assertions
$Q \subseteq M$, the set of questions
$S \subseteq M$, the set of suggestions
$\text{Acc} \subseteq M$, the set of accepted moves

$<$ is a total order on $M$, the order of utterance

$CG$, the common ground, the set of propositions treated as if true by all $i \in I$ at $t$
(This includes propositions about the discourse scoreboard itself.)

$DR$, the set of discourse referents, corresponding to entities entailed to exist in $CG$

$\text{QUD} \subseteq Q \cap \text{Acc}$, the ordered set of questions under discussion at $t$, s.t.
for all $Q \in \text{QUD}$ there is a $g \in G_{com}$ such that $g$ is the goal of answering $Q$, and
for all $Q \in \text{QUD}$, it is not the case that $CG$ entails an answer to $Q$

$CC$, a set of sets of **constitutivity commitments** in effect at $t$, such that
For all $i \in I$, there is a $CC_i$ which is the set of practical interpretive commitments $i$ has at $t$, and

$CC = \{ CC_i \mid i \in I \}$
$CC_{com} = \{ cc \mid \forall i \in I: cc \in CC_i \}$, the set of the interlocutors' common constitutivity commitments at $t$.

---

238 Adapted from Appendix B in Roberts (2012).