The Dissertation Committee for Derek E. Anderson Certifies that this is the approved version of the following dissertation:

SYNTACTICISM AND THE SEMANTIC TURN

Committee:

Sahotra Sarkar, Supervisor

Cory F Juhl

Johan A Kamp

Joshua Dever

Jordi Cat

SYNTACTICISM AND THE SEMANTIC TURN

by

Derek E. Anderson, B.A.; MA

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2016

Acknowledgements

I would like to thank Sahotra Sarkar for sharing his knowledge of the history of logical empiricism, for his insights into the philosophies of Carnap and Neurath, and for all of his help in making this dissertation come together. I would also like to express my appreciation to Cory Juhl for helping me see my way outside of the framework of semantic realism and into a realm of deep confusion. Thanks are also due to Josh Dever for drastically increasing my understanding of logic and to Hans Kamp for his tireless efforts to make me engage more deeply with my subject matter. Special thanks are also due to Briana Schroeder, Jonathan Vanderhoek, Casey Woolwine, Nora Berenstain, and Melissa Kibbe for helping me find my way through this process.

Syntacticism and the Semantic Turn

Derek E. Anderson, PhD

The University of Texas at Austin, 2016

Supervisor: Sahotra Sarkar

There are tensions between empiricism and semantics. The logical empiricists Otto Neurath and Rudolph Carnap were committed to rejecting semantics and treating the logic of science and by extension all of philosophical discourse as a matter of syntax alone, until Tarski convinced Carnap to embrace a semantic theory of truth. Once converted to the semantic paradigm, Carnap attempted to give a criterion of significance that would rule out metaphysics as meaningless while preserving scientific discourse as semantically interpreted. Neurath was convinced from the start that the scientific worldview would be corrupted by a semantic theory of truth. In this dissertation I explore the nature and extent of syntacticism, the view of language as syntax without semantics, and its potential to support a thoroughly empiricist approach to the scientific worldview. I argue that Neurath was right to claim that logical empiricism would not survive the semantic turn in philosophy. I also argue that scientific naturalism about semantics entails Carnap's empiricist criterion of significance; thus, attempts to naturalize semantics are covert extensions of the logical empiricist program.

Table of Contents

Chapter 1: Introduction

Behind closed doors at the Congrès Descartes in Paris, 1937, Neurath and Carnap engaged in their final face-to-face confrontation over the legitimacy of a semantic conception of truth.[1] Neurath, presenting his essay "The Concept of Truth and Empiricism," proposed that an empirical approach to the language of science should limit itself to investigating only syntactic relations between sentences. Semantic relations between language and reality should be rejected, their existence unsupported by empirical evidence. Neurath went further; he claimed that to admit semantics into the scientific worldview would spell the end of empiricism and a return to the bad old days of Aristotelian metaphysics.

Carnap opposed Neurath, following Tarski in maintaining that the semantic conception of truth is correct and unobjectionable and that it cannot be replaced by a fully syntactic approach (Mancosu 2008). Moreover, Carnap regarded it as an open empirical question whether the semantic method would give rise to what the Vienna Circle considered to be metaphysical pseudo-problems. In the conclusion of his presentation he suggested,

> [T]he group of those who have reservations about the semantical concepts...will not carry out public polemics against semantics as a whole until the further developments let transpire, first, whether or not the work in the domain of semantics is fruitful for science and especially for the general task we have set ourselves of an analysis of science and, second, whether or not the feared danger of slipping back into metaphysics is real. Therefore, they will not characterize semantical concepts as a whole as metaphysical but will only criticize single

[1] The details of this exchange are reported in the second-to-last section of Mancosu (2008).

> specific formulations that they might find objectionable especially if they indeed give rise to pseudo problems.[2]

The way was now open for Carnap and those who would follow him to pursue semantic analyses of scientific and philosophical sentences. While Carnap conceded that some semantic analyses might be shown to be objectionable or unscientific, the semantic approach itself was now to be regarded as beyond general criticism.

The application of formal semantics in the service of philosophical analysis has proven to be perhaps the most ubiquitous legacy of the linguistic turn. Contemporary questions about knowledge, causation, truth, essence, mind, thought, belief, necessity, realism, and morality are regularly addressed by invoking methods of formal semantics. The roots of this methodological legacy may be traced back at least as far as Frege, but Carnap's acceptance of Tarski's semantic concept of truth has proved to be an especially crucial step for its widespread adoption within analytic philosophy.

Mancosu (2008) concludes his investigation into the Vienna Circle's debate over semantics by raising the question: is Neurath's proposal a defensible one? Can the language of science be treated in fully syntactic terms, and is this required by empiricism? My investigations suggest that these questions may be answered in the affirmative. I will argue that Carnap's reasons for accepting Tarski's semantic concept of truth were not decisive, and I will argue that Neurath was prophetic in warning that accepting a semantic concept of truth would lead away from empiricism and towards old-fashioned metaphysics.

Neurath and Carnap were not always opposed in their thinking about truth. During his 'syntactic phase' Carnap developed a sophisticated theory of logical syntax

---

[2] From Carnap's typescript, "The semantic conception of truth," (1937), translated by Mancosu (2008).

that elided any semantic notions. This logical syntax would complement Neurath's physicalism by providing a basis for a science of the language of unified science. In order to assess whether Carnap was right to ultimately reject Neurath's proposals in light of Tarski's (1935) definition of truth for formal languages, we must first become clear about the forms of syntacticism that Carnap and Neurath were defending and the extent of their alliance up until 1935. In chapter 2, I discuss Neurath's physicalist syntacticism and its connection with his anti-metaphysical empiricism. In chapter 3, I explore Carnap's syntacticism, including its relation to a formalist epistemology of mathematics and his understanding of Gödel incompleteness. Once syntacticism is clearly articulated, chapter 4 inquires into Carnap's reasons for rejecting it and turning to semantics. I argue that there was nothing inherently wrong with Carnap's syntacticism that forced him to abandon it, nor was Carnap implicitly committed to semantics during his syntactic phase. In chapter 5, I turn to Neurath's reasons for resisting the incorporation of Tarski's semantic theory of truth into the fold of logical empiricism. In chapter 6, I make the case that Neurath's prophetic warning has proved accurate. I argue that semantics is largely responsible for the surge of metaphysical philosophy that characterizes the latter half of the 20th century by investigating several influential metaphysical arguments in which semantic claims play a crucial role.

Chapter 2: Neurath's Physicalist Syntacticism

Syntacticism is not a particular theory but rather a type of theory. It entails a constrained approach to theorizing about language that limits the scope of investigation to syntactic entities and relations that hold between expressions in virtue of their syntactic properties. Semantic entities (such as propositions, contents, and truths), semantic properties (such as being true and being false) and semantic relations (such as reference and satisfaction) are systematically omitted.

The kinds of syntacticism advocated by Neurath and by Carnap during his syntactic phase are importantly different along several dimensions. This chapter will articulate their two versions of syntacticism, exploring their motivations and highlighting similarities and differences.

Neurath's conception of syntacticism flows from his physicalism, the treatment of linguistic entities as physical particulars, empirically observable and located at particular places in space and time. The language of science is a material tool that is constructed for purposes that accord with the goals of society. The effectiveness of a society's language at a given time reflects its development under various historical and material influences, but the language of science can be consciously fashioned into a more effective empirical tool through considered and directed manipulations. We can organize exact rules for producing new sentences on the basis of sentences we have already accepted—in this way, we can make sense of scientific theories themselves and make decisions about which theories to accept—but according to Neurath we should reject any conception of truth as correspondence with reality.

In order to understand the motivation and philosophical situation for Neurath's syntacticism, we must start with a brief excursus into Neurath's political philosophy and its intersection with his epistemology.

2.1 Putting Epistemic Authority in the Hands of the People[3]

Neurath's politics were radical and he fashioned his life so as to be an agent of social progress. Neurath's goal was to facilitate empirically informed social engineering at a massive scale—an engineering project that would be responsive to all aspects of human life including moral justification, economics, and even the language we speak—in the service of liberation from oppression and the construction of a more just and efficient society. This goal grew out of a commitment to socialism and a commitment to the idea that scientific knowledge is a tool of emancipation.

Because the social engineering project was total, and because it needed to be empirically informed in all of its aspects, Neurath saw a pressing need for a unified science, a means of bringing all aspects of scientific knowledge to bear on questions relevant to social decision-making. He recognized that decisions about what we might do as a society interface with scientific knowledge of all scales and domains. In his discussion of sociology in the framework of physicalism, Neurath writes,

> Certainly different kinds of physical laws can be distinguished from each other: for example, chemical, biological, or sociological laws; however, it can *not be said of a prediction of a concrete individual process that it depend on one definite kind of law only*. For example, whether a forest will burn down at a certain location on earth depends as much on the weather as on whether human intervention takes place or not. The intervention, however, can only be predicted if one knows the laws of human behavior. *That is, under certain circumstances, it must be possible to connect all kinds of laws with each other*. Therefore all laws,

[3] My understanding of Neurath's political agenda and its connection to his epistemology owes a large debt to Cartwright, Cat, Fleck, and Uebel (1996), especially the third section written by Cartwright and Cat; many of the themes of that piece are reflected in this section.

> whether chemical, climatological, or sociological, must be conceived as *parts of a system*, namely of *unified science*. (Neurath 1931d/1983, p. 59, emphases in original)

This observation is relevant in any society, even our own. Informed decisions require knowledge drawn from a vast and diverse landscape of scientific knowledge. Should we cut carbon emissions? Should we build a more powerful particle accelerator? Should we immunize ourselves against smallpox? Should we reform the meat industry? How should we construct and modify our laws to reflect the attitudes of the people? Answers to such questions require knowledge of physics, climatology, economics, sociology, immunology, manufacturing, chemistry, and many other areas. Neurath's goal of a unified science was meant to give procedures for making all varieties of scientific law simultaneously applicable in the service of collective decision-making.

As a socialist, Neurath held that the grand social engineering project was to be pursued en masse—not guided by philosopher kings, but by the community of people whom it would serve. For this reason, scientific knowledge must be made accessible to the public. This provides a reason to develop a universally spoken language for unified science that is capable of expressing all scientific knowledge claims. Neurath maintained that such a language must (for inescapable practical reasons) be rooted in common every-day language. The need for a scientifically informed public was also the impetus for Neurath's method of Isotype, a pictorial language designed to help people understand statistics and other scientifically relevant information (Neurath 1974, Chapter 7).

Neurath maintained that humility with regard to human knowledge is an essential component of responsible social action. We must recognize our limitations and not be

swayed by anyone who (aspiring to be a philosopher king) purports to go beyond those limitations. Thus we must be on guard against any form of pseudo-rationalism that might attempt to placate the masses with false promises of certainty. This is a key nexus where Neurath's epistemology meets his politics. In his "Anti-Spengler," Neurath (1921) begins with the following words:

> Politics are action, always built on inadequate survey. But a world-view, too, is action; embracing the manifold universe is an anticipation of unpredictable efforts. In the end all our thinking depends on such inadequacies. We must advance, even without certainty! The only question is whether we are aware of it or not.

Neurath is making two points here. The first is the inescapable necessity of acting without perfect knowledge. Both politics and intellectual life require us to make decisions and move forward with inadequate information. We must do the best we can with the knowledge we have. Second, our thought processes must take our epistemic shortcomings into account. When we are not aware of our epistemic limitations and proceed as if we can always justify our actions, then we are vulnerable to manipulation and deceit by those who are willing to lay claim to certainty through rational insight.

Neurath elaborates this theme in a passage from an earlier piece (Neurath 1913) in which he writes:

> Most of our contemporaries rely on their insight and want to leave the decision in all things to it once and for all. Their starting-point is the view that given enough thought one could at least determine which manner of action has the greater probability of being successful, should certainty be impossible. That there are cases in which one faces several possibilities of action quite helplessly, is denied or declared so highly improbable that no sensible man need give it any further thought. Men of this type are mostly of the opinion that if difficulties turn up,

> sharper thinking will have to lead to the goal; they completely fail to see that even the sharpest thinker can end up with several conclusions of equal value if premises are lacking. Whoever adheres to the belief that he can accomplish everything with his insight, anticipates in a way that complete knowledge of the world that Descartes puts forward as a far-off aim of scientific development. This pseudorationalism leads partly to self-deception, partly to hypocrisy...The pseudorationalists do true rationalism a disservice if they pretend to have adequate insight exactly where strict rationalism excludes it on purely logical grounds.

The notion of 'insight' that Neurath is up in arms against in this passage is a notion of philosophical insight or knowledge through reason alone, the sort of thing that empiricists are opposed to. Complete faith in pure reason is the mark of pseudorationalism. It is the belief that insight can always fill in the gaps in our knowledge that are left open by empirical evidence, and that therefore sufficiently insightful people can know what we should do when the question of what to do is left unanswered by objective empirical evidence. Those who invoke rationality in support of decisive action in all cases purport to know what cannot be known by the masses on the basis of empirical evidence. Thus, through self-deception about their own epistemic limitations, the pseudo-rationalist lays illegitimate claim the role of intellectual leader.

Neurath grants that empirical knowledge does not tell us decisively and univocally what we should do. The solution, however, is not to pretend to have decisive rational insight. For Neurath there is no easy solution to this problem, and it is imperative to resist the temptation of easy solutions:

> Without [God] and with full consciousness of imperfect insight, to act with force and coherence is a difficult task. Whoever attempts this task is disinclined to overestimate insight in any way, as if it were to give wings to our action. As a man he comes to terms with the fact that comprehensive goals are defended on

> insufficient grounds, and indeed they must be so. He abhors self-deception about this...The necessary unity in action is undermined if insight by itself is to bring the final decision. (Neurath 1921)

The attitude recommended is skepticism about rational insight. This attitude is necessary to preserve the necessary unity in action. "Necessary unity in action" is what is required for making decisions as a society. How does an overestimation of insight undermine the necessary unity in action? The problem is that overestimation of insight puts epistemic power into the hands of those few who successfully lobby to be considered insightful in ways that outstrip available evidence.

This is the connecting point between Neurath's socialism and his empiricism. The overestimation of rational insight, together with the metaphysical dictums that such insights produce, is to be opposed by a strict empiricism. The objective is to keep epistemological power in the hands of the people and out of the hands of elite thinkers by grounding knowledge in what is publically observable. Simultaneously, the proper conception of empirical knowledge will be one that makes knowledge available to everyone, or as many as possible. Neurath's physicalism, his advocacy of syntacticism, and his vigilance against metaphysical realism all flow from these socialist-empiricist objectives.

2.2 Neurath's Physicalism: First Aspect

Neurath's physicalism is a thesis about language. He formulates this thesis in a number of places, and differences in his formulations indicate that there are two aspects of physicalism. The first concerns the nature of language as a physical system; the second

concerns the proper composition of the language of unified science in light of the politico-epistemological objectives outlined above. Neurath does not explicitly distinguish these two aspects, indicating that he regards them as closely connected.

In its first aspect, physicalism is a thesis about the proper understanding of the nature of language: language is a physical formation that cannot be treated as different in kind from the rest of the physical world. As such, it has no special properties beyond those of other physical systems. Most notably, the language of science does not express anything beyond its own physical structure. The study of language is the study of physical processes, and that any act of saying is a spatio-temporal arrangement. This spatio-temporal arrangement exhausts what is expressed by language. Neurath (1931a) writes,

> What is scientifically expressible is no richer in fundamental relations than the symbols on a Morse tape which the telegrapher reads as they are sounded by his apparatus. In a sense unified science is physics in its largest aspect, a tissue of laws expressing space-time linkages – let us call it: *Physicalism*. (Neurath 1931a/1984 p.49)

This passage poses a small puzzle. In the first sentence, Neurath gives a statement of his physicalist syntacticism: all that is expressed by a linguistic expression is its own physical structure. A sentence of the language of science only 'expresses' the physical properties of its parts and the relations that obtain between them and other expressions. Immediately after giving this ultra-deflationist picture of scientific language, Neurath goes on to say that the language of unified science expresses all physical space-time linkages! The laws of physics would seem to be much more than the physical relations

between expressions, but somehow Neurath does not agree. The puzzle is resolved when we recognize that Neurath takes the laws of physics to be syntactic strings constructed in a physical medium; thus expressions of the laws of physics express nothing more than the physical properties of and relations between ink patterns on a piece of paper. This points the way toward understanding the connection between physicalism and Neurath's opposition to semantics, which we will return to shortly.

The unity of the physical nature of language—the fact that all of language has a physical structure—is part of what Neurath has in mind when he talks about the unification of science. "All these disciplines [geology, botany, zoology, and the science of science itself] are constructed of the same bricks, as it were." (Neurath 1931a) These bricks are the components of the physical structure of language. Neurath emphasizes the same point again when he writes, "For 'physicalism' it is essential that *one* kind of *order* is the foundation of all laws, whichever science is concerned, geology, chemistry, or sociology." The unity of order Neurath is concerned with is uniformity in the physical nature of the expressions of the language of science.

On Neurath's picture, science is *not* unified in virtue of capturing the unity of reality in the subject of physics. It would be thoroughly confused to think e.g. of Neurath as advocating the claim that all scientifically investigated phenomena are grounded in the properties of microphysical systems, and that this unity of reality is the ground of the unity of science. Physicalism provides the foundation of all laws in the sense that the physicalist language is the physical medium for the expression of every law, each of which is a syntactic object.

Because language itself is a physical structure, the physical language of unified science can engage with it. This is the sense in which language can talk about itself. In explaining the physical nature of language, Neurath (1931b) writes, "scientific language itself is a physical formation whose structure, as physical arrangement (ornament), can be discussed by means of the very same language without contradictions." Language is a physical entity, part of the physical world, and this fact reveals why there is no problem with using language to discuss language. "The statements themselves also form part of other statements as physicalist elements." There is no physical restriction on constructing statements within statements.

2.3 Neurath's Physicalism: Second Aspect

In its second aspect, Neurath's physicalism is a thesis about the proper composition of the language of unified science; viz. that our adaptation of language for the purpose of unified science should start with ordinary common language and transform it so that all the expressions it incorporates can be controlled by empirical evidence—all other expressions are to be regarded as devoid of sense. In explaining this second aspect, Neurath (1931b) writes, "What matters is that all statements contain references to the spatio-temporal order, the order that we know from physics. Therefore this view is to be called 'physicalism'. Unified science contains only physicalist formulations."

According to Neurath's physicalism, the terms that are to be admitted to the language of unified science are not just those terms that belong to the technical jargon of physics, but can include any of the words that are used to describe ordinary objects that

we observe, such as tables and people. These terms do not need to be completely precise. Neurath (1936e) says, "For example, we do not use completely precise terms when we say: 'Man A formulates: in the room was a table perceived by A.' But this kind of formulation as known in everyday language is always needed where predictions are empirically checked by confronting predictions with protocol statements." What is essentially physicalist about the protocol statement in this example is that it contains no terms that cannot be placed under the control of empirical evidence concerning things that are part of the spatio-temporal order—the order that physics investigates.

This notion that expressions can be placed under the control of the senses plays a central role in Neurath's physicalist criterion for inclusion in the language of unified science. Physicalism is meant to rule out statements that depend for their justification on rational insight in part by excluding certain metaphysical expressions from the language altogether. One of Neurath's innovations is to require that the formation of expressions of the language of science (conceived as a physical medium) must be connected in an appropriate way with our senses:

> Our knowledge of phenomena is controlled by sight, hearing, tasting – our sense organs...A statement which cannot be controlled is a thesis devoid of sense. Those who thus succeed in formulating a system of laws which they apply in predicting events were best regarded as "representatives of a scientific conception of the universe". (Neurath 1931a/1984 p. 48)

Neurath is not interested here in providing a theory of empirical justification, but rather in specifying what it is for a system of laws to count as admissible into the system of unified science. The requirement is causal: a necessary condition for incorporating a

statement into the system is that the appearance of that statement be controlled—i.e. manipulated according to our designs—by the sense organs. This is not an epistemological requirement, but rather a requirement on the type of social engineering that is appropriate for the construction of a language for unified science.

This is the point where Neurath's formulation of physicalism comes closest to what Carnap (1963) claims was the Vienna Circle's official doctrine of physicalism in his intellectual autobiography: "The thesis of physicalism, as originally accepted in the Vienna Circle, says roughly: Every concept of the language of science can be explicitly defined in terms of observables; therefore every sentence of the language of science is translatable into a sentence concerning observable properties." It should be noted that Neurath's physicalism as presented here does not entail this strong form of translational reductionism. Specifically, Neurath's program does not entail that every sentence that expresses a scientific law can be translated into an observation sentence.

2.4 Physicalism as Social Program

To invoke an anachronistic metaphor, unified science can be thought of as a kind of computer program. The language of unified science—the code in which the program is written—is specified in physical syntactic terms. That is, the expressions of the language of unified science are identified by their physical properties, but these physical properties are coordinated with syntactic types (for more on this procedure, see the discussion of Carnap's pure vs. descriptive syntax in chapter 3). Expressions of unified science are treated as physical objects without semantic properties; the relevant features of an

expression are exhausted by its physical structure and the relation of that physical structure to the structure of other expressions according to their syntactic classifications. The computations that the program runs are given by transformation rules—statements within the body of unified science that specify inferences that can be made. The purpose for which the computer program is designed is to connect statements generated by the sense organs with predictions generated by the laws of various sciences. These predictions are then used as an aid to the social planning of action. The code must be constructed in a physical medium, since there is nothing else to construct it from, and its functioning must be established through social intervention, since the relevant operations on statements can only be carried out by groups of people within a linguistic community. The community as a whole implements the program of science.

The picture of science-as-computer-program is borne in Neurath's discussion of scientific laws. Neurath (1931d) writes: "Unified science contains all scientific *laws*; these can be connected without exception. Laws are not statements; they are directions for obtaining predictions from observation statements (Schlick)." Here we find two ideas that are connected with the computer program metaphor. Scientific laws—mere syntactic strings—function as parts of inferential chains connecting observation statements with prediction statements. Theoretical statements thus play an important role in the code that generates predictions on the basis of observation, but they do not stand for facts. The idea that all scientific laws "can be connected without exception" means that laws from any set of scientific domains can occur in a single inferential chain.

Neurath's view of scientific laws as inference rules within the program of unified science indicates that the language of unified science can include expressions that are not strictly under the control of the senses. Neurath indicates in many places (Neurath 1915; Neurath 1916; Neurath 1930; Neurath 1934) that he accepts a Duhemian view of science according to which observations can be equally well accounted for by many different incompatible sets of laws. Hence, an expression might be incorporated in the language of physicalism in a role that is not directly responsive to any set of sensory stimulations, but features only in e.g. abstract formalisms that are used in the process of drawing inferences from observation sentences. Being under direct sensory control is not necessary for being an expression in the language of unified science.

The conception of laws as inference rules also sheds light on the earlier puzzle of how Neurath could maintain that laws conceived as mere syntactic strings express space-time linkages: when the strings are part of a system of inference that is implemented within the social fabric of a community, they 'connect' the events that are observed with events that are predicted. An event occurs; it is observed by the sense organs of some individual who, as part of the program of unified science, records the event with an observation sentence. Scientific laws are then used in inferences to prediction sentences, which are used in guiding our interactions with future events.

The idea that all scientific laws can be connected with one another reflects Neurath's (1931b) conviction that "under certain circumstances it must be possible to link the laws of all sciences with each other to make *one* definite prediction." Predictions about a town we plan to build must take account of the geography, the weather, potential

natural hazards, economic factors, sociological factors, and so on. It is out of concern with this interconnectedness that Neurath writes, "For 'physicalism' it is essential that *one* kind of *order* is the foundation of all laws, whichever science is concerned, geology, chemistry, or sociology." (Neurath 1931b/1984 p. 54, emphasis in original) The order Neurath is speaking of is the syntactic medium of scientific prediction; what is required is that the various laws can be made to *physically work together*.

Scientific laws are not the only components of the language of unified science that are used for deriving predictions from observations. Neurath also grants a central place to the implementation of symbolic logic. "The evolution of modern logic makes possible the organization and utilization of all research as upon the foundations of mathematics; and at the same time it clarifies the application of mathematical and other calculi to concrete subject matter. We thus become clearer concerning such matters as the relation between what are called pure and physical geometry, or the logical structure of the probability calculus and its relevance to all the concrete sciences." (Neurath 1937b/1984 p.175) Here, Neurath makes oblique reference to the sorts of logical investigations Carnap (1934) is engaged in, which we will turn to in great detail in chapter 3. Neurath's endorsement of symbolic logic as central to the program of unified science shows the close affinity that he sees between his own work and Carnap's work on logical syntax.

Making statements physically work together is a social engineering project. Statements don't interact on their own; they only interact through the actions of groups of people. Despite making assertions like "laws are not statements," Neurath is not really

giving a linguistic theory, but rather proposing and advocating for a social engineering project that would enforce a pattern of language use within a global community, a pattern that Neurath hopes would facilitate the project of unified science and ultimately the project of providing epistemic power for the masses. One might call the project a 'mechanization of knowledge' in the service of Neurath's global socialist engineering project. In this vein, Neurath writes,

> Everywhere we find a growing sense of technical organization, a sense in harmony with the extension of that new scientific conception of the universe which forges a powerful weapon by the unification of science. No matter in what country or continent they may be, those who regard themselves as simple laborers in solving the riddle of life unconsciously join forces whenever they devote time and effort to the clarification of science and whenever they systematize and interpret with the aid of logic and mathematics all that we perceive through the senses. To predict what will happen and to guide one's actions accordingly is the greatest triumph of earthly striving, the concrete success of human effort which does not make use of *theses devoid of sense* but is rooted in the soil of Physicalism. (Neurath 1931a/1984 p. 51, emphasis in original.)

2.5 Syntacticism: Comparing statements with statements

"An unblemished syntax is the foundation of an unblemished unified science. Language is essential for science; within language all transformations of science take place, not by confrontation of language with a 'world', a totality of 'things' whose variety language is supposed to reflect. An attempt like that would be metaphysics." (Neurath 1931b) The transformations of science are processes that transform statements—physical structures—into other statements. "Thinking in terms of language as physical process is the starting point of all science." But these transformations are specified independently of any relation between the expressions of the language of science and the world apart from

those expressions. The transformations of statements are governed by principles that are fully syntactic, in the sense that they only relate statements with other statements.

Bearing in mind the anachronistic metaphor of unified science as a computer program implemented within the sociolinguistic practice of a community, it is easier to understand what Neurath means when he says that "statements are always compared with statements, certainly not with some 'reality', nor with 'things'[.]" The mechanized system of science proceeds from observation statements to prediction statements. The inputs to the system are statements, the internal procedures are defined over statements, and the outputs are statements. Neurath takes this process to include all that we are licensed to say about scientific practice. To go beyond this picture and speak of a reality that corresponds to the statements of science is to pass into the realm of metaphysics.

When Neurath says that statements are always compared with statements and never with reality, the phrase "compared with" bears some scrutiny. Expressions are not to be compared somehow with reality, as a picture might be, and assessed for truth or falsehood on the grounds of such comparison. Neurath (1934) qualifies this position by pointing out that statements can be compared with other physical objects in the way that physical objects can be compared with other physical objects. For example, we can say that the statement "This chair has four legs" has more words than the chair has legs. We can describe the physical properties of a statement and compare those properties with the physical properties of other systems.

When the impossibility of comparing statements with reality is contrasted with the necessity of comparing statements with other statements, how should we understand

the special way in which sentences can be compared with other sentences? The position being stated is a form of syntacticism: the scientifically relevant relations between statements are given by a comparison of their physical syntactic structure. The scientifically relevant properties of an expression are syntactic, and the syntactic properties of an expression are identified with physical properties.

Neurath's physical syntacticism—the position that statements are only compared with statements—is a weapon of unified science against metaphysics. In the first instance, Neurath rejects the idea that statements can be assessed for truth or falsehood through comparison with reality: "Statements are compared with statements, not with 'experiences', not with a 'world' nor with anything else. All these meaningless duplications belong to a more or less refined metaphysis and are therefore to be rejected." Neurath takes semantic relations themselves to belong to metaphysics, and thus thinks they should be rejected.

But, more than this, Neurath (1934) advocates Carnap's idea that a syntactic approach to language reveals broad swaths of philosophical metaphysics as meaningless: "Precisely for the purpose of evading such idealistic metaphysics, physicalism tries to replace pseudo-content statements (Carnap's 'content language') by statements about language conventions (Carnap's 'formal language')." Sentences that purport to state metaphysical theses are excluded from the language of unified science and replaced with metalinguistic discourse: more statements about statements. This appeal to Carnap's methodology of demonstrating the impossibility of translating sentences of metaphysics into the formal mode (to be discussed below in chapter 3) indicates that Neurath

embraced syntacticism *for the purpose of doing away with metaphysics*, in keeping with his broader anti-rationalism. Hence, one of the goals of physicalism is that it should render metaphysical statements meaningless: "everything that was put forward as philosophy by scholastics, Kantians, phenomenologists, is meaningless except that part of their formulations that can be translated into scientific, that is physicalist, statements." (Neurah 1931b)

The elimination of metaphysics is achieved through the physicalist syntactical treatment of language. "To one who holds the scientific attitude, statements are only a means to prediction; all statements lie in one single plane, and they can be combined, like all parts from a workshop that supplies machine parts. *Physicalism knows no 'depth', everything is on the 'surface'*." (Neurath 1931c /Neurath 1973 p. 326) It is through understanding the expressions of the language of science as components of a mechanized process of knowledge production (or, in our anachronistic conception, as components of a syntactically specified computer program) that Neurath's physicalism deprives philosophy of its claim to access depths of reality that lie below the surface of what is scientifically observable. This anti-metaphysical motive for syntacticism will be important to recall when we turn to the question of Neurath's opposition to Carnap's semantic turn in chapter 5.

2.6 The Encyclopedia, A Public Repository of Knowledge

The output of the program of unified science is a collection of statements, observations and predictions, together with laws and procedures for going forward. Many

of these laws and procedures will be constructed through the use of logical and mathematical techniques. These laws and procedures can be refined through investigation; hence, part of the program's output may be an updated set of instructions about how the program should proceed.

Neurath is opposed to characterizing this output as a 'system' of the kind that he attributes as the goal of science according to Descartes. The purpose of unified science is not "to reach an absolute point from which all particular things should somehow radiate" (Neurath 1936d). The output of the program of unified science is never to be regarded as absolute truth; nor is it to be regarded as complete. Moreover, none of the sentences that compose the output at a given time are to be regarded as certain or as irrevocable. For these reasons (among others) Neurath advocates the choice of the term "encyclopedia" for this output.

An encyclopedia is a physical structure. It is a product of a particular society at a particular place in history. There may be more than one at a time even within a single society, and these multiple encyclopedias need not coincide with one another at all, although they may need to confront one another. This picture is not very different from the actual sociology of science, in which different groups of scientists with competing theories present their data and their competing explanations and try to work things out.

According to Neurath, scientific progress is characterized as the transformation of one encyclopedia into another. "The march of science progresses from encyclopedias to encyclopedias. It is this conception that we call *encyclopedism*." (Neurath 1936d) We

continue to improve and change our encyclopedias as we are confronted with new data, new desires, and new decisions about the directions our societies should take.

Encyclopedism is a deflationary view from the point of view of the systematic metaphysician-philosopher who regards theories as attempts to capture the true nature of reality. This systematic philosopher, the antithesis of Neurath, thinks of the progress of encyclopedias as the evolution of scientific knowledge toward a state of perfection, toward a perfect match with reality. The systematic philosopher's picture is thus similar to the picture of biological evolution according to which organisms evolve toward a state of perfection. Neurath's encyclopedism might be thought of as more analogous to the naturalistic picture of biological evolution according to which the evolution of organisms is not determined by some perfect end state but is instead driven aimlessly by features present in the environment.

2.7 Rejecting "True"

For a number of reasons, Neurath is concerned to restrict the use of the word "true" within the language of unified science. This is not a concern with a theory of truth itself, but rather with the use of the term in public discourse.

One source of uneasiness concerning the word "true" that has been noted by commenters (particularly Mormann 1999) is that Neurath rejects the idealization of a perfect body of scientific knowledge. This rejection is connected with his encyclopedism. He writes, "I propose that one no longer use the term 'the system of science' or any other similar terms, and that one equally avoid all expressions that sound as if they supported

the absolutism of the 'system'. We should never say that certain formulas are 'unshakable', 'definitely free from contradiction', 'absolutely true', nor that they 'approximate' such a state more and more, as if this were something determined or determinable." (Neurath 1936d/1984 p. 145) The fact that there is no ultimate encyclopedia is one reason that we should not speak of "absolute truth" or, what usually comes to the same thing, "truth".

Another reason for Neurath's rejection of the use of the term "true" is that we cannot be certain that any of our statements will not be revised in the future. Neurath's argument against truth concerning certainty is tricky and has confused several philosophers, including Carnap. What does Neurath mean by 'certainty'? He writes, "When we say that one statement is more certain than another, we maintain something regarding our conduct in this respect; for example that we do not intend to spend more time and effort in order to test its truth; moreover, that we do not foresee that the development of science must soon change it, in other words, what would be necessary to do in this case, we do not feel obliged to do." (Neurath 1936d) Certainty is thus construed as a psychological inclination and a physical process, but one that is governed by the program of unified science insofar as that program dictates our conduct regarding how we spend our time and effort and how we update our encyclopedias.

The necessary lack of certainty is connected with the necessary incompleteness of any encyclopedia that we have, and the fact that we can have no justification for refusing absolutely to revise some part of any encyclopedia. Because we can always gather more data, we can never regard an encyclopedia as ultimate. Thus, Neurath's argument about

certainty *does not* turn on an identification of the concept of truth with the concept of certainty, as Carnap would later maintain. It must be understood as an argument concerning the fact that we cannot regard any possible physical repository of knowledge as complete. Neurath extends this point even to logico-mathematical statements, arguing that what we now regard as analytic we might soon regard as contradictory.

Another reason that Neurath regards truth with suspicion is the fact that the language of unified science must, as a matter of historical necessity, be constructed using inexact terms—what Neurath calls *Ballungen*. "The theories with their scientific symbols must lead us in the end to statements that can be checked by mean of statements of ordinary language, for example the following formulations: 'at a temperature of so many degrees even large waterfalls freeze', 'when the interest rate rises, the exchange rate of loans falls', etc. We shall call these expressions and formulations of the common language 'clusters' to distinguish them from scientific formulations." (Neurath 1936e/1984) These clusters, or *Ballungen*, are concepts that cannot be given exact rules of interface with the program of unified science, but are nevertheless essential to that program because they form the basis of the common language with which we must speak even in doing science. To abandon these concepts altogether would be to abandon the connection that history has established between our senses and our observation statements. Since *Ballungen* cannot be given an exact analysis, nor can they be removed from our encyclopedia, the encyclopedia cannot be regarded as an exact or perfect reflection of reality.

None of these considerations, however, gives the deepest reason for Neurath's rejection of truth. The reasons presented thus far are consequences of treating knowledge as a syntactic product of the program of unified science, which follows from Neurath's physicalism. The deepest reason for Neurath's rejection of truth follows from his physicalism itself. In this connection, Neurath writes:

> The study of language can perfectly well be combined with the study of physical processes; for one always stays in the same field. In staying within the closed area of language one can express everything. Thus statements are always compared with statements, certainly not with some 'reality', nor with 'things'...[idealistic and realistic elements] can be completely eliminated if the transition is made to pure unified science...If a statement is made, it is to be confronted with the totality of existing statements. If it agrees with them, it is joined to them; if it does not agree, it is called 'untrue' and rejected; or the existing complex of statements of science is modified so that the new statements can be incorporated...*There can be no other concept of 'truth' for science*. (Neurath 1931b/1984 p. 53)

It is because language is to be treated as a physical medium, and because the use of language is to be specified with regard to its physical syntax only, that the notion of semantic truth should be to be rejected by proponents of Neurath's program of unified science. A strict syntacticism simply does not admit of a semantic theory. Neurath's syntacticism is part of his physicalism, and it is the basis for his other stated reasons for rejecting the semantic notion of truth.

2.8 Redefining "true"

> It has become evident that the use of the terms 'true' and 'false' easily leads to all kinds of difficulties. One can completely renounce the use of these terms, but one can also try to redefine them appropriately. It would, for example, be perfectly expedient to use the term 'true' for all statements that are 'valid' for us in the sense given above, that is, are either part of our encyclopedia or can be deduced

> from it. ‘False’ would then be the qualification of the statements that are in contradiction with the encyclopedia. (Neurath 1936e/1984, p. 161)

Neurath’s thought is that we could continue to use the words “true” and “false” if the use of those terms was properly controlled by the senses—that is, if they were made to be indicators of the relations between sentences and encyclopedias.

Neurath is of two minds about the question at hand, whether we should ban the use of the word “true”, excise it from the language of unified science, or include it but impose strict rules for its use that would render it harmless. Sometimes he warns against the danger to empiricism that the term poses, sometimes he argues that the term should be deployed as a marker indicating that a sentence belongs to the encyclopedia. One gets the impression that he thinks the term is dangerous and should be done away with if possible, but that as a political objective the total elimination of “true” strikes him as unachievable.

On the one hand, the use of the term “true” is part of public discourse. But—as Neurath was keen to point out—the use of the term in public discourse is not strictly in conformity with a semantic definition. Within his circle of academic contacts, however, it was a much more daunting task to get people to stop using the terms “true” and “false”, certainly among his logician friends. In general, especially in today’s philosophical setting, a syntacticist is wise to tread lightly in advocating the rejection of the term altogether.

What should be noted about Neurath’s redefinition of “true” is that it is fully consistent with the physicalist syntacticism of the program of unified science. According to his recommendation, the term would function as a tool in inference—that is, in the

transformation of sentences from one encyclopedia to the next. It would also function as everyday shorthand for indicating which sentences one takes to belong to the encyclopedia at the present time.

2.9 Against Accusations That Physicalism Implies a Coherence Theory of Truth

Neurath is accused of being a coherence theorist by some of his contemporaries, most notably Schlick (1934) and Hempel (1935). These accusations are based on claims that Neurath makes such as the following:

> Each new statement is confronted with the totality of existing statements that have already been harmonized with each other. A statement is called correct if it can be incorporated into this totality. What cannot be incorporated is rejected as incorrect. (Neurath 1931d/1984, p. 66)

> Alongside the present system of statements there is no further *'true' system of statements*. To speak of such, even as a conceptual boundary, does not make any sense. *We can only state that we operate today with the spatio-temporal system suitable for physics*, and that we obtain successful predictions in that way. This system of statements is that of unified science—that is the standpoint we can call physicalism. (Neurath 1931d/1984, p. 61

These statements and others Neurath makes along the same lines should not be taken as an endorsement of a coherence theory of truth. Neurath does not have a theory of truth; he rejects all theories of truth as metaphysics. The program of unified science leaves truth out of account, just as Newtonian physics leaves Aristotle's antiperistasis out of its account of projectile motion, just as the Big Bang theory leaves God out of its account of the origin of the universe.

Neurath (1934) presents a defense of himself against Schlick's (1934) denouncement of physicalism as entailing a coherence theory of truth. Schlick charges Neurath's physicalism with four failings, which amount (in Schlick's estimation) to an endorsement of a coherence theory of truth: (i) physicalism lacks a firm ground of absolute certainty, (ii) it lacks an unambiguous criterion of truth, (iii) it fails to allow for agreement between knowledge and reality, and (iv) it fails to acknowledge that empirical evidence cannot be written down (because empirical evidence is experience, and an experience is not a sentence). Neurath (1934) obliquely responds to each of these points in turn, outlining countervailing considerations raised from within the physicalism picture:

> In opposition to Schlick, I advocate the following tenets:
> (1) All content statements of science, and also those protocol statements that are used for verification, are selected on the basis of decisions and can be altered in principle.
> (2) We call a content statement 'false' if we cannot establish conformity between it and the whole structure of science; we can also reject a protocol statement unless we prefer to alter the structure of science and thus make it into a 'true' statement.
> (3) The verification of certain content statements consists in examining whether they conform to certain protocol statements; therefore we reject the expression that a statement is compared with 'reality', and the more so, since for us 'reality' is replaced by several totalities of statements that are consistent in themselves but not with each other.
> (4) Within radical physicalism statements dealing with 'unsayable', 'unwritable' things and events, prove to be pseudo-statements.

These points must be understood within the scope of Neurath's physicalism. When he says that 'reality' is replaced by several totalities of statements that are consistent in themselves but not with each other, we must recognize that these totalities

are just possible physical structures, potential encyclopedias. Consistency with an encyclopedia is defined over physical syntactic structure. The program of unified science mechanizes the procedure (which is implemented in a group of language users) for verification. Verification is a matter of determining on the basis of input observation statements which statements may be added to the encyclopedia and which statements should be changed or subtracted from the encyclopedia. Neurath grants to Schlick that there is no place for a notion of agreement with reality in this conception of science; his point is to emphasize that this fact is no grounds for labeling him a coherence theorist. He is *not* saying that truth consists in coherence with an encyclopedia.[4]

That physicalism rules out statements dealing with unsayable and unwritable things and events is to say that such sentences are neither under the control of the senses, nor do they play a role in making predictions. Such statements fail to interface with the rest of the program of unified science. As such, they are to be considered pseudo-statements. Hence, Neurath denies any significance to the claim that what counts as evidence is experience itself. What he is interested in denying is that evidence consists in the phenomenology of experience. Against this doctrine, Neurath identifies evidence with the class of sentences that can be controlled directly by the sense organs.[5] The demand

---

4 Neurath's picture is **very** far removed from the sort of coherence theory that was most salient at the time, the sort advocated by philosophers such as Joachim (1906) who contended that truth was identical with conceivability, that conceiving anything required conceiving of all of its relations to other conceptions, thus that only whole possible worlds were conceivable.

5 We might pause to ask: is this identification of evidence with observation sentences a necessary component of syntacticism? I believe the answer is no. Syntacticism requires that evidence not be identified with some semantic thing or property. Hence, it cannot be identified with e.g. knowledge (c.f. Williamson 2000) if knowledge entails truth. Nor can evidence be identified with

for intrinsic justification is replaced by a demand for causal constraints, in the spirit of mechanizing the production of knowledge.

Neurath's first point that statements are admitted into unified science on the basis of decision should not be confused with any kind of conventionalism about truth. The choice of symbols for use within the program of unified science is not a matter of accepting a semantically interpreted linguistic framework (as Carnap, in his semantic phase, will later say). It is purely a question of engineering a physical system to produce desired physical results. Physical structures must be chosen and implemented within the machine. The project of arriving at the correct language of science is continuous and social. As a social process, there are many collective decisions that need to be made. The decision procedure draws on the program of unified science itself. The program of unified science in turn depends on the social practices that it informs.

---

some class of propositions, where propositions are entities that essentially possess truth conditions. But there are other possibilities for a syntacticist. For example, Neurath could have identified evidence with spatially located objects and events (e.g. a bloody knife or the outcome of an experiment) rather than identifying evidence with sentences that are produced through our interactions with such objects and events.

Chapter 3: Carnap's Syntacticism

We turn now to Carnap's (1934) *Logical Syntax of Language* (hereafter LSL), which provides a logical foundation for Neurath's physicalism. LSL is first and foremost a book about logic. Carnap's approach to logic is not intrinsically bound up with any of Neurath's philosophy—a fact that was to become apparent in his semantic turn. Yet, in his syntactic phase, Carnap's work provides a very thorough and important exposition of the power and limits of syntacticism.

At the outset of LSL Carnap writes:

> We shall see that the logical characteristics of sentences (for instance, whether a sentence is analytic, synthetic, or contradictory; whether it is an existential sentence or not; and so on) and the logical relations between them (for instance, whether two sentences contradict one another or are compatible with one another; whether one is logically deducible from the other or not; and so on) are solely dependent upon the syntactical structure of the sentences. In this way, logic will become a part of syntax, provided that the latter is conceived in a sufficiently wide sense and formulated with exactitude. (p. 2)

The extent of Carnap's syntacticism is determined by his conception of the proper domain of logical inquiry, which Carnap takes to encompass the entirety of legitimate philosophy (the legitimate part of philosophy is the fragment that isn't hopelessly doomed to confusion):

> Apart from the questions of the individual sciences, only the questions of the logical analysis of science, of its sentences, terms, concepts, theories, etc., are left as genuine scientific questions. We shall call this complex of questions the *logic of science*...[O]nce philosophy is purified of all unscientific elements, only the logic of science remains...For this reason we prefer to say: *the logic of science takes the place of the inextricable tangle of problems which is known as philosophy*. (p. 279, emphasis in original)

Carnap's view simply stated: "the formal method, if carried far enough, embraces all logical problems" and thus, "all problems of the current logic of science, as soon as they are exactly formulated, are seen to be syntactical problems." (p. 282) Logic is to be reduced to syntax and because all legitimate philosophical problems are problems in the logic of science, all legitimate philosophical problems are properly resolved through attention to logical syntax.

In his intellectual autobiography, Carnap (1963) identifies four interests that motivated the development of his view in LSL. The first was to "show that the concepts of the theory of formal deductive logic, e.g. provability, derivability from given premises, logical independence, etc., are purely syntactical concepts, and that therefore their definitions can be formulated in logical syntax, since these concepts depend merely on the forms of the sentences, not on their meanings." The second was to "show that many philosophical controversies actually concern the question of whether a particular language form should be used, say, for the language of mathematics or of science." The third, connected with the second, was to "show that everyone is free to choose the rules of his language and thereby his logic in any way he wishes." Carnap called this the *principle of tolerance*. Fourth, because philosophical problems were seen as problems concerned with language, Carnap believed that "these problems should be formulated, not in the object language, but in the metalanguage. Therefore it seemed to [Carnap] that the development of a suitable metalanguage would essentially contribute toward greater clarity in the formulation of philosophical problems and greater fruitfulness in their

discussions."[6] Neurath and Carnap were united in these four motives for giving a fully syntactic account of language (especially the first three).

3.1 Pure syntax vs. descriptive syntax

Yet Carnap's picture is quite distinct from the physicalism that motivates Neurath's syntacticism. Neurath took syntactic entities to be physical particulars. But Carnap identifies a domain of *pure* syntax that is distinct from any physical phenomenon. Carnap argues for the non-physical nature of pure syntax in two different ways. First, he points out that the choice of physical symbols adopted for use in a particular language is irrelevant to that language's formal properties. The designs used by printers, or the fact that these designs are composed of ink, are unimportant for the study of logical syntax. Carnap goes so far as to say that any series of physical things could equally well serve as expressions in a language: "The *syntax* of a language, or of any other calculus, is concerned, in general, with *the structures of possible serial orders* (of a definite kind) *of any elements whatsoever*." Carnap elaborates with an explanation of pure syntax that identifies it as a form of geometry:

> *Pure syntax* is concerned with the possible arrangements, without reference either to the nature of the things which constitute the various elements, or to the question as to which of the possible arrangements of these elements are anywhere actually realized (that is to say, with the possible forms of sentences, without regard either to the design of the words of which the sentences are composed, or to whether any of the sentences exist on paper somewhere in the world). In pure syntax only definitions are formulated and the consequences of such definitions developed. Pure syntax is thus wholly analytic, and is nothing more than combinatorial analysis, or, in other words, the geometry of finite, discrete, serial structures of a particular kind. (p. 6-7)

[6] Quotations and paraphrases are from pp. 54-55 of Carnap (1963).

Pure syntax is distinguished from *descriptive* or *physical* syntax. Descriptive syntax is a distinct type of theory that investigates the syntactic properties of physical systems. Carnap maintains that the relation between pure and descriptive syntax is closely analogous to the relationship between pure mathematical geometry and physical geometry. Carnap (1922) gives an in-depth account of this latter relationship. Geometers are free to investigate any number of distinct geometries within the realm of pure mathematics; meanwhile, inquiry into the nature of physical space requires both experimental evidence and correlative definitions that identify components of some chosen abstract geometry with components of a physical system. Likewise, pure syntax may be concerned with any number of distinct formal languages; meanwhile, inquiry into the nature of a given physical symbol system requires both experimental evidence and correlative definitions, which identify physical properties with syntactic properties.

Carnap (1934) elaborates on this analogy in §25. He mentions Hilbert's axioms as an example of a set of axioms of pure geometry. His formulation of Language I serves as an example of a set of axioms of pure syntax. Carnap gives an example of a correlative definition for physical geometry: "Physical objects of such and such a kind (for instance, light-rays in a vacuum or stretched strings) are to be considered as straight segments." He compares this with a correlative definition for identifying the numeral zero in a given physical syntax: "'Nu(x)' is to be taken as true when and only when a written character having the figure of an upright ellipse ('0') is to be found at the position x." In both cases, the correlative definition is assumed as an axiom—a part of the logical machinery—of

the physical theory. The identification of syntactic types with physical types, then, is itself an aspect of one's choice of logical language. For Carnap, this is an immediate consequence of identifying pure syntax as a part of geometry, given his view of geometry and its relation to physical structure.

Carnap establishes the abstractness of syntax in a second way, namely through the arithmetization of syntax induced through Gödel numbering. Each term of the syntax language is assigned a unique number. Rules are then established for assigning unique series-numbers for sentences (sequences of terms). A series-number is the product of a sequence of powers of the first *n* primes >2, where *n* is the number of terms in the sentence. The *j*th prime in the sequence is raised to the power of the term-number of the *j*th term in the sentence. This establishes a unique series-number for each sentence. Similar rules are implemented for assigning unique series-series-numbers for proofs (sequences of sentences). The whole procedure is done in such a way that claims about syntactic properties of expressions are guaranteed to be materially equivalent to sentences about arithmetic properties of numbers. Once the systematic equivalence of syntactic sentences and arithmetical sentences is established, Carnap concludes: "All the sentences of pure syntax follow from these arithmetical definitions and are thus analytic sentences of elementary arithmetic." (p. 57).

Carnap says that the most important reason for the arithmetization of syntax is that it enables him to talk about what is 'possible' within a syntactic theory without introducing modal notions such as possibility into the theory itself. Since his theorems are equivalent to theorems about the properties of numbers, they will hold in any domain in

which arithmetic holds. Yet this equivalence also establishes the abstractness of pure syntax, since syntax is part of arithmetic and arithmetic is abstract.

This is consistent with holding that syntax is part of geometry, since geometry too can be construed as a collection of theorems about the arithmetical properties of sequences of numbers. Indeed, it is clear that Carnap regards the arithmetization of syntax as very closely analogous to the arithmetization of geometry. In §25, he puts the following definitions side by side and gives them the same roman numerals (emphasis in original):

> I. Arithmetical geometry
>
> *A partial domain of arithmetic* which (in the usual method of arithmetization, namely by means of co-ordinates) is concerned with ordered triads of real numbers, the linear equations occurring between them, and the like.
>
> I. Arithmetical (or pure) syntax
>
> *A partial domain of arithmetic* which (in the method of arithmetization previously explained) is concerned with certain products of certain powers of prime numbers, the relations between such products, and so on.

Where Neurath maintains that language should be regarded as a physical structure, Carnap's approach to syntacticism assumes a very different starting point, one that identifies syntax as a part of arithmetic. So I pause to emphasize that Neurath was still very much in support of Carnap's project. Neurath (1930) writes: "The modern scientific world-conception owes its successes partly to the new symbolism that can be used for the purification of language...[I]t seems to be true to say that our most fruitful symbolism for mastering the concrete had to have very non-concrete antecedents." Rather than seeing a dramatic divide between his own form of syntacticism and Carnap's,

Neurath views Carnap's starting point—his concern with pure mathematics—as a necessary antecedent for developing an adequate science of the concrete, physical language of unified science.

3.2 Tolerance and logical pluralism

One thesis that follows naturally from Carnap's pluralist understanding of geometry, together with his identification of syntax as a part of geometry, is his famous principle of tolerance. The principle of tolerance says, roughly, that there are no criteria for acceptability of the axioms of a pure syntactic theory.

Uncharacteristically, Carnap's formulation of the principle of tolerance is vague and metaphorical—no precise definition is offered. Carnap's formulation is written as follows. "*Principle of Tolerance: It is not our business to set up prohibitions, but to arrive at conventions*." Carnap elucidates the principle only a little further, writing:

> *In logic, there are no morals*. Everyone is at liberty to build up his own logic, i.e. his own form of language, as he wishes. All that is required of him is that, if he wishes to discuss it, he must state his methods clearly, and give syntactical rules instead of philosophical arguments. (p. 52, emphasis in original.)

If all the possible axioms of a logical syntax are equivalent to statements about arithmetic, what does pure syntax have to do with arriving at conventions? And what does it mean to build up a form of language if we are not physically building something?

These questions are best answered by attending to Carnap's identification of syntax with geometry. Classically, it was believed that geometry was accurately

characterized by a single set of axioms, namely Euclid's. With the advent of non-Euclidean geometries, the idea that there was a single set of true geometric axioms was replaced with a new conception of geometry according to which various sets of axioms could be investigated independently of one another, with no question of truth arising between them.[7] Carnap's principle of tolerance is an extension of this idea to the domain of logical syntax. Indeed, given the fact that pure syntax is identified as a part of geometry, the principle of tolerance in syntax may be regarded as a consequence of geometrical pluralism.

Bearing in mind the close analogy between pure syntax and geometry, how should we understand Carnap's claim that our 'business' is to 'arrive at conventions'? Given that every theory of pure syntax is a fragment of arithmetic, such theories are only conventional insofar as arithmetic itself is conventional. Physical or descriptive syntax, on the other hand, is partially a matter of convention, since the correlative laws that establish the syntactic properties of physical objects are conventional in the sense that their inclusion in a theory is not dictated by evidence but rather by the decision of the theorist. Is Carnap's principle of tolerance concerned, then, with descriptive syntax rather than pure syntax?

I think the answer is no. In the case of geometry, the choice of axioms (say, between a Euclidean and a non-Euclidean system) is a choice between distinct pure geometries. There is a question of convention that arises concerning which aspects of a physical system should be identified with aspects of a given geometry within a physical

---

[7] For a very in-depth historical overview, see Nagel (1939).

theory, but this conventionalist aspect only comes into play after we have accepted a principle of tolerance concerning pure geometry. We first accept that there are many equally legitimate abstract geometries, and then we are faced with the question of which geometry is best suited to model the physical world. Similarly, for Carnap, the question of tolerance must arise first in the domain of pure syntax. Conventional decisions about what physical language we should use or about how we might describe a physical system through the application of a pure syntactic theory or about how we might set up a system of physical symbols to reflect the properties of a pure syntax (as when we devise a computer program) are separate from the question of whether there is one true set of axioms for logical syntax. Carnap's principle of tolerance is only concerned with this last question. It is a denial that there is one true absolute set of axioms for logical syntax. 'Convention' is therefore something of a red herring. What is at issue is the admissibility of sets of axioms and inference rules characterizing distinct theories of pure syntax. For Carnap, logic itself imposes no constraint on the form that a logical syntax can take.

To 'build up a language' in the domain of pure syntax is to select a set of rules that determine, "in the first place, the conditions under which an expression can be said to belong to a certain category of expressions; and, in the second place, under what conditions the transformation of one or more expressions into another or others may be allowed." (p. 4) When one gives a set of these rules, they specify or identify a pure syntactic theory. The 'conventionality' involved in such a specification, insofar as there is any, can be no more than an agreement to think about or presuppose a particular set of axioms. That is what Carnap means when he says: "All that is required of [a person

proposing a logical syntax] is that, if he wishes to discuss it, he must state his methods clearly, and give syntactical rules instead of philosophical arguments."

Carnap puts the principle of tolerance to work by developing two distinct syntax languages, Language I and Language II. Language I is constructed to reflect the mathematical constraints imposed by a constructivist methodology. It is finitary and admits only limited quantification. Language II allows the formulation of all of classical mathematics. By the principle of tolerance, neither language is 'more true' than the other. Carnap takes himself to resolve the tension between intuitionist and classical mathematics by showing that these two languages are each acceptable. The difference between intuitionism and classical mathematics is only a difference in choice of language. This serves as a model of the way in which syntacticism renders philosophical disputes less 'deep' than they might seem to be.

3.3 Pluralism and pragmatism

Carnap makes no pretension to have constructed the only two useful languages. Rather, he presents Language I and II as models or examples of the general procedure of building up a logic. How far does the pluralism engendered by Carnap's principle of tolerance extend? Sarkar (2013) argues that Carnap is committed to allowing even inconsistent languages. That is to say, the principle of tolerance implies that there are no logical constraints whatsoever imposed on our choice of logical syntax. For all Carnap (1934) says, we may adopt an inconsistent language if we wished to. Sarkar regards this as an objection that needs to be responded to, and cites Beth (1963) and Gödel (1995) as

raising the objection against Carnap. But for a proponent of the principle of tolerance, there really is no reason to believe that logic itself imposes sanctions against the use of inconsistent languages, especially once it is recognized that inconsistent languages can be safeguarded against explosion. In principle we may discover reasons for adopting inconsistent languages for certain purposes.

Practical reasons are always relevant when constructing a language or choosing to apply it to some domain. On this point, Carnap was influenced by Neurath. In his intellectual biography, Carnap (1963) describes a dispute in methodology between Neurath on the one side and Carnap and Schlick on the other. The dispute starts when Neurath defends materialism against idealism on the grounds that "during the last hundred years, materialism was usually connected with progressive ideas in political and social matters, while idealism was associated with reactionary attitudes." (Carnap 1963, p. 50) Carnap and Schlick maintain that materialism and idealism are pseudo-theses, each as much as the other. Moreover, they initially regard Neurath's methodology as confused, since it presents a historical sociological observation in defense of a metaphysical position. But, as Carnap writes:

> When I suggested that we should not discuss the theses of idealism and materialism but rather the problem of the choice of a language, Neurath accepted this point but tried to turn my weapon against me. The choice of a language form is a practical decision, he argued, just as the choice of a route for a railroad or that of a constitution for a government. He emphasized that all practical decisions are interconnected and should therefore be made from the point of view of a general goal. The decisive criterion would be how well a certain language form, or a railroad, or a constitution, could be expected to serve the community which intended to use it. His emphasis on the interdependence of all decisions, including those in theoretical fields, and his warning against isolating the deliberation of

> any practical question, even that of the choice of a language form, made a strong impression upon my own thinking and that of my friends. (p. 50)

It was under Neurath's influence, then, that Carnap came to accept that practical reason is central to questions of the construction and use of language. The quoted passage reveals two points of contact between Carnap's pragmatism concerning logical syntax and Neurath's social engineering program. First, Carnap concedes that the choice of language is to be made on the basis of practical reasons. In instituting a language for a given purpose, we must know what this purpose is and we must ask whether the language will be suitable for this purpose. Second, Carnap concedes that Neurath is right about the holistic nature of practical decision-making in the social sphere. The purpose for which a language is to be put must be understood in relation to all of the goals of a community. To accept the "inter-dependence of all decisions," and thus the need to align all decisions with the goals of a community, is to endorse a very radically socialist way of thinking about logic, decision-making, and life in general.

Moreover, Carnap (1963) identifies this practical dimension of language choice as a consequence of the principle of tolerance, with implications for the understanding of metaphysical disputes:

> As a consequence [of the principle of tolerance], the discussion of [philosophical controversies] need only concern first, the syntactical properties of the various forms of language, and second, **practical reasons** for preferring one or the other form for given purposes. In this way, assertions that a particular language is the correct language or represents the correct logic such as often occurred in earlier discussions, are eliminated, and traditional ontological problems, in contradistinction to the logical or syntactical ones, for example, problems concerning "the essence of number", are entirely abolished. (p. 54, emphasis added)

Our logical understanding—our capacity for describing and distinguishing the syntactical properties that codify different forms of language—gives us the domain of languages we have access to and from which we can select a language form for a given purpose. That is the extent to which logic gives us knowledge; it cannot inform us about the nature of things or their essences. Only practical reasons govern our choice of language forms for particular purposes. There is no further question about which language is correct, so there is no further question of whether one description or another captures the essence of reality as it is. Our thinking about issues in metaphysics is determined completely by our choice of syntax, which is governed by practical reason. When this thesis is combined with the social engineering conception of practical reason endorsed by Carnap in the previous quotation, the result is striking: our metaphysical commitments should be whatever best serves the goals of our community!

Yet we must not be confused into thinking that metaphysical commitments that come from syntactic entailments have much significance, since the commitments themselves are just meaningless strings, to be regarded as nothing more than artifacts of our choice of language form. Only confusion follows when this is forgotten. For this reason Neurath advocates against allowing idle metaphysical expressions into the physicalist language by requiring that all expressions either play some role in inferring predictions or be placed under the direct control of the senses. For a strict syntacticist, cognitively significant statements of empirical facts cannot be distinguished from senseless statements of metaphysics on the grounds that the former are literally

meaningful while the latter are not, because meaningfulness is part of semantics. Hence, Neurath's endorsement of something like an empiricist criterion of significance is purely a matter of social engineering—constructing a language for the purposes of unified science.

The reasons for constructing a language are not issued from the domain of logic. From a Carnapian perspective, giving a logical syntax is very much like developing a computer program. You are free to explore and implement a system of symbols and a system of rules governing the evolution of those symbols for whatever purpose you see fit, be it in the realm of mathematics or the realm of science. The fact that Carnap places absolutely no logical constraints on a theory of logic is thus very much aligned with Neurath's physicalist syntacticism. For Neurath, the constraints on the physicalist language are dictated not by some transcendent assessment of logic but by the needs and goals of the society that employs the language.

3.4 Validity and Logical Consequence

The logical pluralism engendered by the principle of tolerance also reflects a deep and important aspect of syntacticism: the rejection of any transcendent criterion of validity or logical consequence. From a contemporary perspective that embraces a model-theoretic semantics, validity is characterized in terms of truth preservation in all models: an argument is valid iff any model that satisfies the set of premises $\sum$ also satisfies the conclusion C. If the notion of semantic truth is excluded from logic—as it is when logic is characterized exclusively in terms of syntax—how are we to understand the notion of

validity? There is no longer any theory-transcendent characterization; validity can only be specified syntactically. Validity itself is thus subject to the principle of tolerance. Logic provides no in-principle constraints on what may be counted as a valid argument.

Carnap is explicit about this point after proving that the principle of complete induction and the principle of selection (i.e. the axiom of choice) are analytic in Language II. Carnap acknowledges that in proving their analyticity, he presupposes that the principles themselves are available in the syntax language. Consequently, he writes:

> The proofs of Theorems 1 and 2 must not be interpreted as though by means of them it were proved that the Principle of Induction and the Principle of Selection were materially true. They only show that our definition of 'analytic' effects on this point what it is intended to effect, namely, the characterization of a sentence as analytic if, in material interpretation, it is regarded as logically valid. The question as to whether the Principle of Selection should be admitted into the whole of the language of science (including also all syntactical investigations) as logically valid or not is not decided thereby. That is a matter of choice, as are all questions concerning the language-form which is to be chosen. (cf. the Principle of Tolerance...) (p. 124)

Carnap thus clearly understands that questions of validity are matters of choice when building a language. As stated in the principle of tolerance, there are no constraints placed on such choices by logic itself. What can be regarded as valid is only constrained by whatever practical reasons determine the language chosen.

One purpose that occupies Carnap (1934) is something he considers to be a chief task of the logical foundation of mathematics: "to set up a formal criterion of validity, that is, to state the necessary and sufficient conditions which a sentence must fulfill in order to be valid (correct, true) in the sense understood in classical mathematics." (p. 98) What had once been hoped for was a syntactic criterion of validity such that "the question

of its fulfillment or non-fulfillment could in every individual instance be decided in a finite number of steps by means of a strictly established method." (p. 98-99) Failing that, the hope would be to give a "method of derivation" by which we specify (by some finite means, either directly or schematically specify) a set of primitive axioms and a set of rules of inference that would together suffice to imply every truth of classical mathematics.

Carnap takes Gödel's incompleteness results to dash both of these hopes. Nevertheless, Carnap maintains that a complete criterion of validity for mathematics can be obtained—consistent with Gödel's results—through the application of what he terms the *method of consequence*: "a method of deduction which depends upon indefinite individual steps, and in which the number of the premises need not be finite." (p. 100) According to Carnap, the notion of consequence he develops for Language II provides a complete criterion of mathematical validity—it determines the complete set of mathematical truths. Carnap proves this by showing that every logical sentence of the system is L-determinate: either analytic or contradictory. Since the notions of analyticity and contradiction are fully syntactic, the proof (if correct) shows that Gödel incompleteness need not be understood in semantic terms.

3.5 Syntacticism, Incompleteness, and the Analytic Indefinite

Gödel incompleteness entails that for any system of axioms S sufficient to express Robinson arithmetic—for example, Carnap's Language II and any other language

capable of expressing all of classical mathematics[8]—some numerical properties are not definable in S.[9] Equivalently, some classically valid arithmetical sentences are not decidable in S. Hence there are valid arithmetical sentences that are not decidable from within Language II. Yet it is Carnap's aim to establish a complete criterion of mathematical validity within Language II in terms of analyticity. Thus, to accommodate Gödel incompleteness, analyticity must be defined such that the undecidable sentences of arithmetic are nevertheless analytic. Most obviously, Carnap must give a syntactic criterion of analyticity that entails that the Gödel sentence for Language II is analytic.

Whether a sentence is analytic, contradictory, or synthetic is determined through the application of a series of syntactic transformation rules for replacing sentences with other sentences. The application of the sequence of rules to a sentence G always produces one of three results: (i) the result is a sentence "0=0" in which case the original sentence G is analytic; (ii) the result is a sentence "0≠0" in which case G is contradictory; (iii) the sequence of rules is exhausted and neither "0=0" or "0≠0" results, in which case S1 is synthetic. The rules are guaranteed to terminate after a finite number of applications, so the definition of analytic is definite in this sense. Some of the rules, however—specifically, the rules of valuation and evaluation—are not definite in the sense that applying them involves making inferences from countably infinite sets of sentences.

The sequence begins with nine rules of reduction.

---

8 Language I is also strong enough to express Robinson arithmetic.

9 A numerical property P is definable in S in the relevant sense iff there is a formula A(x) of S such that (i) if a number n has P then S proves A(n) and (ii) if n does not have P then S proves not-A(n).

(1) Definitions are eliminated.

(2) The resulting sentence is converted into conjunctive standard form.

(3) Redundant conjuncts and disjuncts are eliminated. Then "0=0" is substituted for any disjunctions of S and ~S, and "0≠0" is substituted for any conjunctions of S and ~S. Then any instance of "0≠0" that is disjoint with another sentence is eliminated[10] and any instance of "0=0" that is conjoined with another sentence is eliminated. Then any sentence that is conjoined with "0≠0" is eliminated. These sub-rules are repeated as many times as they are applicable before moving on to rule 4.

(4) Limited existential operators are replaced with negated limited universal operators.

(5) Where t_1 and t_2 are substitutional variables that range over numerals (in Language II every numeral is a "0" followed by some number of accent marks): Any instance of "$t_1 = t_1$" is replaced with "0=0". Any instance of "successor(t_1)=successor(t_2)" is replaced with "$t_1=t_2$". Any instance of "0=successor(t_1)" is replaced with "0≠0." In effect, whenever you have an equation, you subtract one from both sides until you reach "0=0" or, if you reach "0=successor(t_1)", you replace the resulting sentence with "0≠0".

(6) This rule concerns the elimination of sentential variables. For a sentence G with free sentential variables, let *f* be the first free sentential variable of G. Then G is replaced with a conjunction of new sentences G_1 and G_2, substituting "0=0" for *f* in G_1 and "0≠0" for *f* in G_2. For example, "*f*→*f*" would be replaced with "(0=0→0=0) &

---

[10] This rule is only applied once per disjunction, so that in the case of "0≠0 or 0≠0" only one disjunct is eliminated; what results is a single contradiction "0≠0".

$(0\neq0\rightarrow0\neq0)$". For a sentence G with bound sentential variables, let *f* be the first sentential variable of G bound by the existential operator (the universal operator is treated in the same way as the free variable). Then G is similarly replaced with a disjunction of new sentences G_1 and G_2 where "0=0" is substituted for *f* in G_1 and "0≠0" for *f* in G_2. For example, "$\exists f_1 \exists f_2 (f_1 \rightarrow f_2)$" would (after two applications of the rule) become "$(0=0\rightarrow0=0)$ or $(0\neq0\rightarrow0=0)$ or $(0\neq0\rightarrow0\neq0)$ or $(0=0\rightarrow0\neq0)$". Hence, this rule is guaranteed to transform tautologies (or contradictions) invoking sentential variables into other tautologies (or contradictions) of a format that can effectively interact with the other rules of reduction.

(7) This rule eliminates K-operators from the target sentence G. K-operators are numerical operators that return the smallest number meeting some definition, or if there is no such number the value of the operator is 0. For each K-operator, the target sentence is replaced with a disjunction that functions somewhat like a Russellian definite description picking out the smallest number with the numerical property F defined by the K-operator in question if any number has the property or zero if not. The first disjunct handles the case in which there is no number with the property F; in this disjunct the sentence G is replaced with a sentence in which 0 is substituted for the K-operator conjoined with a negative existential sentence asserting that there is no smallest number with the defined property. The second disjunct gives the case in which there exists a unique smallest number *n* with property F; in this disjunct the sentence G is replaced with a sentence in which *n* is substituted for the K-operator.

(8) This rule eliminates limited universal operators from the target sentence G. Such a sentence has the form ∀x(n)(S), where S is a sentence and n is a number

specifying the upper bound of the scope of the universal operator. There are three cases. In the first case, x is not free in S: G is replaced with S (i.e. the limited universal operator is simply chopped off from the front of G if its variable does not occur in S). In the second case, n=0: again, G is replaced with S (again the operator is chopped off because the limitation of the operator is total, i.e. nothing is 'quantified over'). In the third case, n is the successor of some number: then G is replaced with a conjunction of S with n substituted for x together with the sentence $\forall x(n-1)(S)$. This third case is repeated until the second case occurs, effectively producing a conjunction of n copies of S such that each number from zero to n is substituted for x in exactly one copy—this effectively produces a sentence for each value the variable can take. Once the proper case has been dealt with, a sentence is added that is equivalent to the conjunction of $\forall x(j)(S)$ for all $j \leq n$.

(9) After rules 1-8 have been completed, only unlimited operators remain in the target sentence G. Rule 9 slides all of these unlimited operators to the front of G, replacing G with an equivalent sentence written in standard form. Operators that correspond to no free variables are discarded. Negation signs are 'pushed in' to the right of all quantifiers, transposing the quantifiers they cross over in the usual way.

After application of the rules of reduction, another sequence of rules is applied. These are the two rules of valuation and two rules of evaluation. After each rule of evaluation is applied, the procedure must loop back through the rules of reduction. If a non-reduced sentence results from the application of a rule of evaluation, that sentence is reduced before returning to the rules of evaluation.

The first rule of valuation defines the notion of a *valuation* and gives a recursive specification of a hierarchy of types of valuation. A valuation of type 0 is an accented expression. A valuation of type t_1, t_2, ... t_n is an ordered n-ad of valuations of types t_1, t_2, ... t_n.[11] So a valuation may be a series of numerals, or a series of series of numerals, or a series of classes of numerals, and so on. A valuation of type (t_1) is a class of valuations of type t_1. This may be a class of numerals, or a class of classes of numerals, or a class of series, or a class of valuations of type $(t_1{:}t_2)$, and so on. A valuation of type $(t_1{:}t_2)$ is a function from t_1 to t_2, where t_1 and t_2 may be valuations of any type, including n-ads or classes of valuations.[12] These four types of valuation can be combined in all the obvious ways, keeping in mind that the variables t_1, t_2, etc. can be replaced by any type of valuation. So, for example, t_1 may be of type (0,0:0); in that case, the valuation of (t_1) will be of type ((0,0:0)) and the valuation of t_1, t_1, t_1 will be of type (0,0:0), (0,0:0), (0,0:0).

Note that every valuation is constructed from accented expressions. The valuation of a first-order predicate is a class of accented expressions; the valuation of a second-order predicate is a class of classes of expressions; the valuation of a first order functor is an accented expression; the valuation of a second order functor is a class of accented expressions. These syntactic objects will be used to define functions from expressions to

---

11 The numbers appearing in the subscripts of this series only indicate distinctions between variables that range over types; they do not indicate levels in a hierarchy of types.

12 Carnap (1934, p 147) indicates that real numbers may be defined as expressions of the type (0:0), that is, as functions from the class of accented expressions to the class of accented expressions. For every such expression there is a corresponding valuation. Hence, there are uncountably many valuations. Uncountably many of these valuations, corresponding to the uncomputable irrationals (those with infinite Komolgorov complexity), will be undecidable in the sense that there will be no definite procedure for enumerating the series of mappings from accented expressions to accented expressions.

expressions—the rules of evaluation—that will replace every partial expression involving a predicate or a functor with either “0=0” or “0≠0”.

The second rule of valuation constrains the set of valuations that may be chosen for each partial expression in a given reduced sentence without operators. Every accented expression is assigned itself as a valuation. Every series of expressions must be assigned an ordered n-ad composed of the valuations of the components of the series. For any expression that has function-argument form, that expression must be assigned a valuation determined by the valuation of the function applied to the valuation of the argument. These rules guarantee that every expression type is assigned a valuation of the same type.

This second rule does not determine a unique valuation for every partial expression in a given reduced sentence without operators. It only constrains the type of valuation that can be assigned. For example, a predicate can have as its valuation any class of accented expressions, but nothing else. A functor of type (0,0:0) can be assigned any function from pairs of accented expressions to accented expressions, but nothing else. The only expressions that have their valuations determined uniquely by the second rule are the accented expressions themselves.

Finally, the two rules of evaluation are applied. The first rule states that if a partial sentence G has predicate(argument) form then if the valuation of the argument expression belongs to the valuation of the predicate, then G is replaced with “0=0”; if the valuation of the argument does not belong to the valuation of the predicate, G is replaced with “0≠0”. The second rule governs the replacement of equations. If a partial sentence G has

the form $E_1 = E_2$ then G is replaced by "0=0" if the valuation of E_1 is identical to the valuation of E_2 and "0≠0" otherwise.

Carnap emphasizes that in every case a valuation is a syntactic property or, equivalently, an equivalence relation between expressions of the language. The valuation of a first-order predicate is a function from accented expressions to sentences, the range of which is the set {"0=0", "0≠0"}. The syntactic property had by the class of accented expressions that map to "0=0" when entered as arguments to this function is merely that they form an equivalence class defined by this function. More generally, every valuation is a function from syntactically determined objects to syntactically determined objects that defines a syntactic property, and every valuation thereby specifies a syntactic property or equivalence class in the domain of its function. So, in the simplest case, the syntactic property that is shared amongst the accented expressions in a given valuation is nothing more than being those expressions that are mapped to "0=0" by the first rule of evaluation when concatenated with a predicate assigned that valuation.

Recall that syntactic properties are materially equivalent to numerical properties. Statements about syntactic objects and relations can be expressed as statements about products of powers of prime numbers. So, for example, the fact that a given accented expression e is in the valuation of a predicate P is equivalent to the fact that there is a function that maps the series-number of "P(e)" to the series number of "0=0". Generally, this function will not be computable. Carnap notes, "As a result of Gödel's researches it is certain...that for every arithmetical system *there are numerical properties which are not definable*." (p. 106, emphasis in original) Hence, informed by Gödel's

incompleteness result, Carnap allows that there may be syntactical properties—for example, valuations of predicates—that are not definable. "For just as for every language there are numerical properties which are not definable in it, so there are also syntactical properties which are not definable in [it]." (p. 114) This is a key feature of Language II, and any language that could possibly be used to express all of classical mathematics.

We call a syntactic property *definite* if there is a recursive procedure for determining all and only those expressions that have that property and *indefinite* otherwise. Carnap writes, "Our attitude towards the question of indefinite terms conforms to the principle of tolerance; in constructing a language we can either exclude such terms (as we have done in Language I) or admit them (as in Language II). It is a matter to be decided by convention." (p. 165) The fact that we are considering a language with indefinite syntactic properties is a matter of choice; it is solely because Carnap's goal is to provide a language in which all mathematically valid statements are analytic.

Using the rules of reduction, valuation, and evaluation, Carnap defines "analytic" and "contradictory" for sentences of Language II and for classes of such sentences. The definition of an analytic class of sentences depends on the definition of analyticity, which in turn depends on a logically preliminary notion: 'analytic in respect of a valuation.' (I present the series of definitions in the reverse order that Carnap does—I start with the definition of analyticity for atomic sentences and work up to the definition for classes of sentences, rather than beginning with classes and working 'backward' as Carnap does—but I preserve Carnap's original numbering of the rules.)

The preliminary notion 'analytic in respect of a valuation B_1' is defined for atomic sentences with no operators—these are the most basic cases, sentences consisting of predicate(argument) or function(argument) form. B_1 is a series of at least two valuations; this is necessary because an atomic sentence contains minimally two expressions that need valuations in order to be subject to a rule of evaluation (i.e. a predicate and an argument or a function and an argument); moreover, an atomic sentence could contain multiple nested predicates and/or functors, e.g. F(A(x,y:z)), so B_1 may have more than two entries.

Carnap defines this notion for three cases. The first two cases tell you how to ignore an operator:

(**DA3.A**) if G is a sentence with form (v)(S) then G is analytic with respect to B_1 if for every valuation B_2 of v, S is analytic with respect of both B_1 and B_2.

(**DA3.B**) If G is a sentence of form (∃ v)(S) then G is analytic with respect to B_1 if for one valuation B_2 of v, S is analytic with respect of both B_1 and B_2.

The third case tells how to evaluate the analyticity of an atomic sentence with no operators. This is where the rules of evaluation take root.

(**DA3.C**) If G has no operators, then G is analytic with respect to B_1 if applying the rules of evaluation on the basis of B_1 to G transforms G into "0=0". G is contradictory with respect to B_1 if applying the rules of evaluation to G on the basis of B_1 transforms G into "0≠0".

At the next level of logical complexity, Carnap defines analyticity and contradiction for sentences more generally. This again involves several cases.

(**DA2.A**) Any sentence G that is not reduced is analytic if the sentence derived from the rules of reduction applied to G is "0=0" and contradictory if the resulting sentence is "0≠0".

(**DA2.B**) If G is reduced and open, then it is analytic if its universal closure is analytic and contradictory if its universal closure is contradictory (i.e. refer to the next two cases).

(**DA2.C**) If G is reduced, closed, and *logical* (i.e. contains only logical terms, no descriptive terms) then:

(**a**) G is analytic if G has the form of (v)(S), where v is a variable of any type and S is a sentence, and S is analytic in respect of every admissible[13] valuation of v; G is contradictory if there is one valuation of v that makes S contradictory;

(**b**) G is analytic if G has the form (∃ v)(S) and S is analytic in respect of at least one valuation; G is contradictory if S is contradictory in respect of every valuation;

(**c**) G is analytic if it has the form "0=0" and contradictory if it has the form "0≠0". If G is reduced, closed, and *descriptive* (i.e. it contains at least one

---

[13] Since the type of valuation is constrained by rule two given the type of variable, the set of all admissible valuations will be different from variable-type to variable-type. If the variable is of the same type as accented expressions then the class of every admissible valuation includes all and only the accented expressions themselves; if the variable is a first-order predicate variable, then every admissible valuation includes all and only classes of accented expressions; and so on up the hierarchy of types. This applies to every variable discussed in the definition of analyticity, but henceforth I will leave out the word "admissible" and assume this is understood.

descriptive expression)[14] then G is analytic if the singleton class containing G is analytic (this is defined according to rule DA1.C below).

Lastly, we define "analyticity" for classes of sentences. A class of sentences R is defined as analytic according to the following rules:

(**DA1.A**) If not all sentences of R are reduced, then R is analytic if the class of sentences that results from applying the rules of reduction to every sentence of R is analytic. R is contradictory if the resultant class is contradictory.

(**DA1.B**) If all sentences of R are reduces and logical, then R is analytic if every sentence of R is analytic. R is contradictory if one sentence of R is contradictory.

(**DA1.C.**) If the sentences of R are reduced and at least one of them is descriptive, then:

(a) If an open sentence occurs in R, then R is analytic if the class of sentences consisting of the universal closure of every sentence in R is analytic.

(b) If the sentences of R are all closed, then the class is analytic if a different class made up of suitable replacements is analytic; a suitable replacement for a sentence S with a descriptive predicate is a logical sentence generated by uniformly replacing

[14] Descriptive expressions are distinguished by convention: descriptive expressions are coded with a distinct set of term numbers. They are also treated differently according to the rules of inference. Other than these, there are no distinctive syntactic features of descriptive expressions. This will be relevant in section 3 below.

> every descriptive expression in S with a logical expression of the same type.

Consequence is then defined with regard to (possibly infinite) classes of sentences R. A sentence S is a consequence of R iff ~S and R is contradictory. Thus, a sentence is analytic just incase it's negation is incompatible with some set of analytic sentences.

Carnap points out that the question of the analyticity of a given sentence or class of sentences always refers in a finite number of steps back to one of the two rules DA3.C and DA2.C.c. But because the syntactic relations encoded by valuations are indefinite, the predicates "analytic" and "contradictory" in Language II are indefinite as well. There is no general method for determining whether a given sentence is analytic, contradictory, or synthetic. (p. 112-113)

We are now in a position to see how Carnap's syntacticism accommodates Gödel incompleteness. Rather than directly proving that the Gödel sentence itself is analytic, Carnap proves a generalization of Gödel's incompleteness theorem and then shows that a number of 'irresoluble' or indefinite sentences are analytic, including the Gödel sentence for Language II. The generalization of Gödel's incompleteness theorem is Carnap's fixed-point lemma.[15] This lemma states that, for any syntactic property P that could be

[15] The significance of this result is emphasized by Sarkar (1992), to which the following exegesis owes a large debt. Sarkar calls the fixed-point lemma "the most interesting new result for language II" that Carnap proves. The details of Carnap's proof were first presented in a similar form in Noonan (1990).

possessed by an expression of Language II, we can construct a sentence G of Language II such that G is analytic iff G has P.

Carnap proves the lemma by walking us through a generalized procedure by which we can construct, for any syntactic property, a sentence that 'means' that it has the property in question. We start with a sentence G_1 consisting of the chosen syntactic predicate P and a free variable x.

G_1 looks like this: "P(x)".

We assume that the variable x has term number n. Let G_2 be a new sentence derived from G_1 by substituting "subst[x,n,str(x)]" for x.

G_2 looks like this: "P(subst[x,n,str(x)]".

This sentence involves of a pair of numerical operators: str(x) is a numerical operator whose value is the series-number of whatever expression is substituted for x. Subst[x,y,z] is a numerical operator whose value is the series-number of the expression that results when the expression with series number y occurring inside the expression whose series number is x is replaced by the expression whose series number is z. (In slightly less careful words: subst[xyz] is a computable procedure that takes the expression numbered by x, then finds expression y inside of that expression and replaces it with expression z, and then returns the series-number of the resulting expression.) Now, let b be the series-number of G_2 and consider the series number that is computed by the function: subst[b,n,str(b)]. This function returns the series number of a sentence derived from G_2 by substituting its own series-number b for x. Call this sentence G_3.

G_3 looks like this: "P(subst[b,n,str(b)])".

The series-number of G_3 = subst[b,n,str(b)] = j, hence $G_3 \Leftrightarrow P(j)$. Since j is the series-number of G_3, we can see that G_3 predicates P of itself.

Since P was arbitrary, this proves that for any predicate, a 'self-attributing' sentence can be produced. Using this lemma, Carnap derives the Gödel sentence for Language II by substituting "~BewSatzII(r,x)" for "P". Call this sentence G_G. ~BewSatzII(r,x) means that in Language II there is no number r that is the series-series number of a proof that terminates with a sentence having the series-number x. In other words, the sentence with series-number x is not provable by Language II. Let g be the series-number of G_G. Then G_G looks like this: "~BewSatzII(r,g)". So G_G 'asserts' its own unprovability.

Carnap shows that G_G is analytic on the assumption that Langauge II is not contradictory. If there were a series-series number that represented a proof of G_G, then Language II would prove BewSatzII(r,g). But since G_G is the sentence "~BewSatzII(r,g)", this means Language II would prove ~BewSatzII(r,g) if it proved BewSatzII(r,g). Thus we know that ~BewSatzII(r,g) is not demonstrable by Langauge II, provided Language II is consistent.

Carnap also proves that Language II is consistent. The proof is long but essentially it consists in the following two-premise argument: (i) Language II proves only analytic sentences; (ii) "0≠0" is not analytic. Therefore, Language II does not prove "0≠0" and so Language II is consistent (since an inconsistent language proves anything). Carnap is careful to point out that this result does not violate Gödel's second incompleteness theorem, since the proof relies on resources such as the indefinite term

'analytic' that outstrips anything that can be defined inside of Language II. Thus, G_G is shown to be both analytic and indefinite.

I pause to emphasize how different Carnap's conception is from the received semantic view of Gödel incompleteness. According to the semantic view, Gödel showed that the set of mathematical truths cannot be captured within a single axiom system. Carnap's understanding of incompleteness is very deflationary compared to this picture. Incompleteness is not characterized against a background of mathematical truths. Rather, it is characterized by the fact that a sufficiently complex syntax can produce sentences within a language that cannot be proved within that language, but which can be derived within stronger languages. (In fact, the Gödel sentence of any language can be simply added to that language as a primitive axiom to create a new, slightly stronger language in which that sentence is definite. There will be a new Gödel sentence for the stronger language, but this sentence will be definite in a language that is slightly stronger still. And so on for every Gödel sentence.)

3.6 Logicism, Formalism, and Mathematical Knowledge

Carnap's aim in searching for a complete criterion of mathematical validity was to provide an epistemology of mathematics inspired by Frege's logicism that incorporated the insights of Hilbert's formalism (Friedman 1999). On Carnap's picture, our knowledge of mathematics is analytic as Frege believed. Mathematical knowledge flows from the preconditions on coherent thinking about anything at all. But pace Frege, Carnap denies that there is one, true logic. Hence he denies that there is one true set of preconditions on

coherent thought. The principle of tolerance reflects the central theme of formalism: one is free to choose any set of axioms when building their language.

But does this picture really fit together, given Carnap's exposition of analyticity discussed in the previous section? A worry arises when we reflect on the relationship between analyticity and mathematical knowledge. Let us take for our example Goldbach's Conjecture. Assuming that Carnap's criterion of mathematical validity is indeed complete, Goldbach's Conjecture is either analytic or its negation is analytic. Yet we do not know whether Goldbach's Conjecture is valid in Language II or not. Moreover, it is possible that there is no proof of Goldbach's Conjecture or its negation. How could a mathematical sentence be analytic yet unknowable? And doesn't this result undermine the epistemological project of logicism?

The root of this puzzle lies in the definition of analyticity, specifically at the point where valuations are introduced. Consider the Goldbach predicate "x is greater than two and x is the sum of two primes". This expression is assigned a valuation V, which is a class of accented expressions. For every sentence S generated by substituting an accented expression E for x, S is either analytic or contradictory. E is a member of V or it is not. If it is, then S is replaced by "0=0" according to the rules of evaluation; if it is not, then S is replaced by "0≠0". But the valuation of "x is greater than two and x is the sum of two primes" is unknown—if we knew what it was, then we would know whether or not Goldbach's Conjecture is valid. If we do not know the valuation of this predicate, how could we have determined it when building Language II? And if it was not determined by our choice of language, then where did it come from?

Perhaps the numbers themselves determined the valuation of the Goldbach predicate. If we gave a semantic theory of Language II, we might maintain that the accented expressions pick out the numbers themselves and the predicate "x is greater than two and x is the sum of two primes" designates a property of numbers. Then we could maintain that V is determined by the numbers themselves together with the semantic facts about the language: an accented expression E belongs to the valuation of the Goldbach predicate iff the number that E designates has the property designated by the Goldbach predicate. More generally, with a semantic theory we could say that mathematical objects and properties determine whether or not a mathematical statement was analytic.

Carnap rejects semantic explanations of where valuations come from. Responding to the charge that the definition of "analytic" is semantic in this way, Carnap writes:

> Do we not by this means arrive at a Platonic absolutism of ideas, that is, at the conception that the totality of all properties, which is non-denumerable and therefore can never be exhausted by definitions, is something which subsists in itself, independent of all construction and definition? From our point of view, this metaphysical conception...is definitely excluded. We have here absolutely nothing to do with the metaphysical question as to whether properties exist in themselves or whether they are created by definition. (p. 114)

Carnap is pointedly opposed to even raising the question of whether properties that exist in themselves explain the valuation of the predicates of Language II. Indeed, any such explanation must be ruled out from the syntactic point of view. But the question remains, how can a valuation be determined by our choice of language yet remain unknown?

Carnap points out that for some sentences that cannot (at present) be conclusively judged, such as Goldbach's Conjecture, "we know under what conditions we should say that the answer had been found." (p. 161) For the Goldbach predicate, we have a method of deciding whether a given accented expression E belongs to its valuation. This is because we have a computable procedure for deciding whether "x is greater than two" is analytic when E is substituted for x, and also a computable procedure for deciding whether "x is the sum of two primes" is analytic when E ix substituted for x. These conditions are spelled out by other definitions of Language II.

This points us toward a deeper puzzle concerning the valuations of predicates that are highly indefinite. Goldbach's Conjecture is a Pi_1 statement and so occupies a relatively low place on the arithmetic hierarchy. We can specify a counterexample to Goldbach's conjecture using definitions formulated in Language II. But other terms such as "analytic" are indefinite to a much higher degree. Carnap proves that a definition of "analytic in Language II" cannot be formulated in Language II. Yet "x is mathematically valid" and "x is an analytic mathematical sentence" must have the same valuation in Language II. How is this valuation determined? More generally, how are the valuations of highly indeterminate expressions determined? And what about statements for which we do not presently have any idea of how we would prove or disprove them, such as the generalized continuum hypothesis?

In considering the admissibility of highly indeterminate expressions, Carnap appeals to the principle of tolerance. He writes:

> The proper way of framing the question is not "Are indefinite (or impredicative) symbols admissible?" for, since there are no morals in logic, what meaning can 'admissible' have here? The problem can only be expressed this way: "How shall we construct a particular language? Shall we admit symbols of this kind or not? And what are the consequences of either procedure?" It is therefore a question of choosing a form of language—that is, of the establishment of rules of syntax and of the investigation of the consequences of these...In any case, the material reasons so far brought forward for the rejection either of indefinite or of impredicative terms are not sound. We are at liberty to admit or reject such definitions without giving any reason. But if we wish to justify either procedure, we must first exhibit its formal consequences. (p. 164-165)

The most striking thought expressed in the previous quotation is that we can accept a definition—for example, Carnap's definition of the indefinite term "analytic"—without any reason or justification at all. More generally, we can give rules of syntax (choose a language) without giving any justification for those rules. Of course, Carnap cannot do so while aiming to reconcile a logicist epistemology with a formalist conception of meta-mathematics, but what Carnap says here indicates that he takes his own goals for developing a logical syntax to be inessential to the nature of logic itself. Giving a Carnapian logicist theory of knowledge for classical mathematics against the backdrop of the principle of tolerance is not a matter of discovering something about the nature of logic that justifies our mathematical claims.

What does it mean, then, to say that *justifying* the admission of indefinite symbols requires exhibiting the formal consequences of that procedure? It is to say that justifications must be offered in the terms of logical syntax. While it is not possible to justify the admission of every term of a language within that very language, according to Carnap it is always possible to justify it using some other language. Hence, Carnap writes

> For every term which is stated in any unambiguous way in a word-language, there exists a formal definition in an appropriate language. Every arithmetical sentence G which is, for instance, irresoluble in the language S_1 is yet determinate in S_1; in the first place there exists a richer syntax language S_2, within which the proof either that G is analytic or that G is contradictory can be stated; and secondly, there exists an object-language S_3 of which S_1 is a proper sub-language, such that G is resoluble in S_3. But there exists neither a language in which all arithmetical terms can be defined nor one in which all arithmetical sentences are resoluble...In other words, *everything mathematical can be formalized, but mathematics cannot be exhausted by one system; it requires an infinite series of ever richer languages*. (p. 222, emphasis in original)

If Carnap is right, then mathematical knowledge claims can always be justified by formal considerations. The definition of a term, or the assignment of a valuation in accord with a given definition of analyticity, must be provided if that definition or that valuation is to be justified, but in principle such things need not be justified—they may simply be tolerated. Recalling the previous quotation, exhibiting the formal consequences is necessary to justification: "if we wish to **justify** [the inclusion of an indefinite expression], we **must** first exhibit its formal consequences." Hence formalization is both necessary and sufficient to justify the inclusion of an indefinite term. Insofar as we have mathematical knowledge, this knowledge is always justified by formal considerations, and formal considerations are always given in the terms of pure syntax.

This may or may not accord with actual mathematical practice. Insofar as mathematicians are willing to accept claims on the basis of intuition without formal explication, they will disagree with Carnap. Thus, Carnap is issuing a substantive epistemological imperative: only regard a mathematical sentence as justified if the definitions and rules of inference used to derive that sentence have been formally specified in some well-defined syntax language or other. There are no morals in logic, but

there are morals in epistemology. Carnap's Fregean motivations for developing a logical syntax of language reside in this imperative. Thus, pace Creath (1990), Carnap's logicism is not "logicism virtually in name only...[having] been absorbed into formalism". Carnap's logicism subsists in the epistemological imperative of formal explication.

I will consider one more puzzle concerning Carnap's logicist epistemology. If pure syntax is a fragment of arithmetic, how can formalization in pure syntax be the grounds of our mathematical knowledge, including our knowledge of arithmetic? If our knowledge of mathematics flows from our grasp of logical syntax, it seems that our epistemic access to logical syntax must be distinct and prior to our epistemic access to mathematics. But then how can logical syntax be a mere fragment of one branch of mathematics?

One place to gain some insight into this conundrum is to return to our earlier considerations about what it means to build a logic or to choose a syntax language. We broached the idea that all of the possible axiom systems we might adopt 'exist' prior to any choices we might make, in the sense that all the numerical properties of products of powers of prime numbers are what they are independently of human conventions. Once we select a syntax language, if our language is strong enough to formulate expressions of arithmetic, then the syntactic rules of our language can be encoded as materially equivalent to sets of arithmetical statements. With a theory of arithmetic in hand, we can analyze syntax in arithmetical terms.

Given the principle of tolerance, we are free to adopt any language we choose. If we were to adopt the finitary Language I, our theory of arithmetic becomes finitary and

restricted in the way that intuitionists restrict it. If we adopt Language II, we accept arithmetic in its classical form. Thus, our theory of arithmetic depends on our choice of syntax language. This dependence is even clearer when we consider what would happen if we were to adopt an inconsistent syntax language. Then our theory of arithmetic would be likewise inconsistent, since we may prove any 'arithmetical theorems' whatever from our inconsistent syntax language. Moreover, while we can prove that the axioms of any logical syntax are equivalent to a fragment of arithmetic, the proof itself relies on assuming some background logical syntax or other to give us the relevant rules of inference. These considerations strongly support the idea that our conception of arithmetic depends on what syntax language we are working with.

Does this show that arithmetic itself, and mathematics more broadly construed, is conventional? It's not clear, because the sense of 'convention' at issue is very unclear. We cannot think of such a convention as a matter of deciding which system of pure syntax to endorse from among a number of antecedently given axiom systems. Nor can we think of the convention as a matter of choosing to use physical symbols in some particular way—Carnap is very clear about the priority of pure syntax over descriptive syntax. Our decisions as physical organisms to make certain physical marks cannot determine anything about pure syntax. This latter point highlights a misleading aspect of the received view that Carnap was a conventionalist about logic in the sense that we somehow create logic by adopting some set of communal practices. Assuming that we are physical beings, the idea that our actions and activities have some bearing on the nature of logic cannot be right on Carnap's account. The image of a person creating or building

a pure syntax language is a metaphor. There is no clear sense in which Carnap's view entails that mathematics depends on convention conceived of as human action.

I think these ruminations reveal that the question of logicism can only arise for Carnap against a background of logical syntax that is strong enough to generate some mathematical theory. When we have a coherent syntax of a certain kind, we can use that syntax to investigate mathematical structures; we can then use those mathematical structures to investigate logical syntax. The situation is reminiscent of Carnap's conception of the development of physical measurement (Carnap 1974). All physical measurements are conducted with physical objects; our measurement practices are refined through improving our understanding of our physical measuring devices, a process that involves using physical objects to make measurements on other physical objects. Likewise, mathematical investigations are conducted with mathematical and syntactic objects; by using formal and mathematical tools, we refine our understanding of formalism and mathematics. Carnap rejects the notion that there is an absolute form of measurement in the domain of the physical. We should understand his theory of mathematical knowledge in a similar light.

3.7 Neurath and Carnap in Alliance

It is clear that there are significant differences between the syntacticism advocated by Neurath and that developed by Carnap. Carnap's investigations are very abstract; in contrast, Neurath's philosophical theories are only interested in what can be located in space and time, in what can influence and be influenced by a group of people. In what

ways are these views allied? I conclude this chapter with a brief consideration of the nature of the alliance between Neurath and Carnap during Carnap's syntactic phase.

Neurath plied Carnap to produce a theory of the logical syntax of language to serve as the basis for a science of unified science. He expresses his optimism that Carnap would carry out such a project when he writes (1931b):

> If one wants to speak of the further development of physicalism, it can probably be expected that the attempt that Carnap undertook in his Logisher Aufbau der Welt (1928) will be repeated in order to create the syntax for unified science in the sense of physicalism as represented here. The work on unified science replaces all former philosophy.

Neurath required a science that was capable of characterizing the program of unified science. To this end, his project relied on the development of a theory of logical syntax that was sufficient to meet his purposes. Carnap's logical syntax of language is well suited for Neurath. The cogency of Carnap's project is thus of great significance to Neurath's physicalism.

On the other hand, Carnap's project would seem to rely not at all on Neurath's politico-epistemological physicalism. Carnap was quite content to allow his theories to remain fully abstract. Although he was clearly influenced by Neurath, and at least for a while his logical goals were in line with Neurath's agenda, there was nothing in Neurath's program of unified science that was essential to Carnap's concerns.

Chapter 4: Carnap's Semantic Turn

At the end of his recollections on LSL in his intellectual autobiography, Carnap (1963) writes

> A few years after the publication of the book, I recognized that one of its main theses was formulated too narrowly. I had said that the problems of philosophy or of the philosophy of science are merely syntactical problems; I should have said in a more general way that these problems are metatheoretical problems. The narrower formulation is historically explained by the fact that the syntactical aspect of language had been the first to be investigated by exact means by Frege, Hilbert, the Polish Logicians, and in my book. Later we saw that the metatheory must also include semantics and pragmatics; therefore the realm of philosophy must likewise be conceived as comprising these fields.

Why did Carnap abandon syntacticism and embrace the idea that philosophy must be investigated using semantics? In this chapter I argue that he was not rationally compelled to embrace Tarski's semantic theory of truth by some failing of his earlier theory of logical syntax. Nor did he become convinced that his own theory was really a semantic theory in disguise. Rather, Carnap's change of direction is to be explained solely in terms of his guiding principle that what is accepted by mathematicians and scientists may be embraced by empirically-minded philosophers. This principle, together with Carnap's impression that Tarski had "investigated by exact means" the field of semantics from a mathematical perspective, explains why he was willing to take a new turn.

A number of authors have offered competing stories in an attempt to understand Carnap's reasons for first rejecting a semantic theory of truth and then accepting it in the face of bitter protestations from Neurath. I will defend my interpretation of events

through an examination of three of these competing stories, presented by Coffa (1987), Creath (1990), and Oberdan (1992). A central theme in all of these pieces is that Carnap's theory of logical syntax was implicitly committed to semantics; I will rebut this claim where it arises.

4.1 Coffa on Carnap and Tarski

Coffa (1987) investigates the role that Tarski's theory of truth, and more generally Tarski himself, had in moving Carnap to accept a semantic theory of truth. According to Coffa, Carnap (1934) had already embraced a semantic theory of truth, but had not recognized it as such. His interaction with Tarski brought him out of his syntactic dogmatism and into the light of realization that his own work was really shot through with semantics.

Coffa claims that Carnap's starting point for his work on a criterion of mathematical validity in LSL was his recognition, in light of Gödel's theorems, that "truth and theoremhood are different things, and that consequence and proof are equally far apart." Coffa goes on to exclaim that Carnap never seems to have seriously doubted that, in the range of mathematics, proof was one thing and truth an entirely different one. Coffa claims that, "If urged to present his explicandum in the dreaded material mode Carnap would have said that the languages under consideration in LSL are interpreted as dealing with the natural numbers...and all classes that may be built up from them" and moreover that every sentence of the syntax language is either true or false.

It should be noted that Carnap does not characterize the situation this way. Carnap writes only of valid mathematical sentences, not mathematical truths. In fact, Carnap goes to remarkable lengths to keep the notion of truth out of his explication of the sense in which mathematical validity outclasses provability. We are therefore licensed to suspect that Coffa is attempting to stack the deck in favor of a semantic understanding of Carnap's system; hence, Coffa's remarks about the semantic nature of Carnap's explicandum should be stricken from the record.

According to Coffa, the appearance of semantics within Carnap's account of mathematical validity occurs in Carnap's definition of analyticity for Language II. Coffa writes, "The essentially semantic nature of this new approach becomes clear when we notice that the centerpiece of the new characterization of truth is not a generalization of the notion of inference, but the radically new idea of *valuation*." The notion of valuation is semantic, according to Coffa.

Coffa rightly identifies Carnap's reason for implementing the strategy of valuation: Gödel's results prove that inference rules represented within Language II will never be sufficient to decide every valid mathematical sentence. Strangely, Coffa says that Carnap himself provides no explanation for his change of strategy. This is patently not the case. In his explanation for supplementing the definition of analyticity with rules of evaluation, Carnap writes:

> In the case of a predicate- or functor- variable, however, the [method of inference by rule of substitution] does not succeed; a fact which has been pointed out by Gödel...As a result of Gödel's researches it is certain, for instance, that for every arithmetical system there are *numerical properties which are not definable*...Thus,

> in the case of a [predicate], we cannot refer to substitutions but must proceed in a different way. (Carnap 1934, p. 106-107.)

Carnap then goes on to give the rules of valuation and evaluation. Carnap's motivation is clear. The introduction of the notion of a valuation is explicitly and self-consciously motivated by considerations of Gödel incompleteness.

To make the case that a valuation is a semantic entity, Coffa turns to Tarski's investigation into a notion of mathematical truth that could capture what was left out of an incomplete system. Coffa observes that Tarski's search for a definition of mathematical truth led him to separate two notions of definability. The first is syntactic: a predicate P is definable iff any string containing P can be proved equivalent to another string that does not contain P. This is a notion of definition as elimination. The second notion of definability is model-theoretic: we start with a "well defined semantic object (usually an individual or a class) and we ask whether there is in a given language, a propositional function whose extension is the object in question." Coffa goes on to say "the problem of model-theoretic definability does not concern itself with the introduction or elimination of expressions but, rather, with the question of whether the expressive power of a language suffices to capture a certain semantical object."

According to Coffa, the semantic desiderata of the model-theoretic notion of definition is reflected in Tarski's convention (T): a theory of truth for language L must entail that for every sentence "S" of L, "S" is true iff S. Coffa writes, "The point of convention (T) is that its satisfaction will guarantee that the truth-predicate ['is true'] will have the class of true sentences as its extension and thus qualify—in the semantic

sense—as a definition of it." Thus, the adequacy condition that Tarski gives for a definition of truth *presupposes* the existence of a class of true sentences. An adequate definition of "is true" must 'capture' this class. So the semantic notion of definability presupposes the existence of semantic objects and properties. Understood in this way, the attempt to define truth for mathematics begins with the truths of mathematics as a class unto itself. Then we attempt to find a definition of "truth" that captures the set of mathematical truths by giving a propositional function whose extension is exactly that set. Thus, a semantic definition of truth along the lines that Tarski pursues allows us to characterize Gödel incompleteness in its familiar semantic form: for any set of axioms, there will be truths of arithmetic that cannot be proved from that set.

We should pause to note how different this conception of Gödel incompleteness is from the one that Carnap engages with. Carnap proves the fixed-point lemma that for every syntactic predicate P, a sentence S can be constructed in which P takes the series-number of S itself as an argument. Given that 'not-provable in Language II' is a decidable predicate in Language II, it can be proved that there is a sentence G_G of Language II that predicates 'not-provable in Language II' of its own series number. The supposition that there is a series-series number that represents a proof for this sentence G_G leads to contradiction; hence, we infer that the sentence G_G itself is valid (as long as Language II is not contradictory), but there is no proof in Language II of G_G. This exhausts what has been discovered by Gödel. There is no assumption about an independent class of truths of mathematics. Nor is Carnap even concerned with the prospect of defining truth for mathematics.

Tarski's (1936) definition of truth appeals to satisfaction of a sentential function by a sequence of objects. Coffa says that Tarski's notion of satisfaction is identical with Carnap's notion of valuation in every way except for the fact that Tarski's notion is not constrained to syntactic objects. Coffa thus attempts to sell the idea that very little in the way of philosophical significance separates the two notions, and that therefore Carnap's definition of analyticity was, for all philosophical purposes, a semantic definition. This is very much like saying that a syntactic theory of language is a semantic theory of language in every way except for the fact that semantic theories are not constrained to syntax. What is added to a syntactic theory to make it a semantic theory is very philosophically significant. Likewise, what would need to be added to Carnap's notion of valuation to make it equivalent to Tarski's semantic notion of satisfaction is very philosophically significant.

A valuation is a function from syntactic expressions to syntactic expressions. Given the arithmetization of syntax, a valuation is equivalent to a function from numbers to numbers. These functions are not necessarily computable. Nevertheless, they are simply relations between abstract objects (numbers, strings, or what have you). They are not relations between the syntax language and the world. Their introduction does not presuppose any domain of language-independent truths. In this way, they are metaphysically inert and epistemically independent of knowledge of the world, in just the way that a syntactic theory of language in general is broadly metaphysically inert and epistemically independent of knowledge of the world.

By Coffa's own lights, the difference between a semantic (model-theoretic) approach to truth and a syntactic approach is very philosophically significant. In the semantic case, one starts with a class of semantic objects and properties as given. In the syntactic case, one can only appeal to relations between syntactic items and definitions amount to nothing more than eliminations of terms. The deep significance of the distinction between Tarski's semantic satisfaction relations and Carnap's syntactic valuations should have leapt out at Coffa.

Having supposedly established the close connection between Carnap's notion of valuation and Tarski's notion of satisfaction—a connection which should in fact be rejected—Coffa goes on to wonder why Tarski "got there first" (referring to a semantic theory of truth). From Coffa's perspective, Carnap almost discovered a semantic theory of truth before Tarski did, and might have done so if he hadn't been hung up on verificationism.

Coffa notes that Carnap gives an argument against defining a notion of truth within logical syntax, but he calls it "one of the worst Carnap ever advanced." Coffa quotes Carnap's argument as follows:

> Truth and falsehood are not proper syntactical properties; whether a sentence is true or false cannot be gathered in general from its design, that is to say, by the kinds and serial order of its symbols. (This fact has usually been overlooked by logicians, because, for the most part, they have been dealing not with descriptive but only with logical languages, and in relations to these, certainly, 'true' and 'false' coincide with 'analytic' and 'contradictory' respectively, and are thus syntactical terms.) (p. 216 in Carnap 1934)

Coffa claims that this argument is indefensible and also (shockingly) claims that its conclusion is false. "Truth can be defined in what [Carnap] called the syntax of a language," Coffa tells us. He makes this claim on the grounds that "Carnap's syntax-languages were indistinguishable from Tarski's semantic metalanguages and truth was consequently definable in them." Here Coffa's confusion about the supposed semantic nature of Carnap's logical syntax is coming to the fore. If Carnap had accepted Tarski's model-theoretic account of definition by assuming the existence of semantic properties, then presumably he could have granted that truth was definable in the way that Tarski attempted. But Carnap was explicitly committed to remaining in the domain of syntax—that is Carnap's most central constraint on a theory of language. If a separate domain of truths is not presupposed—if we really limit ourselves to the syntactic domain—then it simply does not make sense to try to define truth within logical syntax.

In fact, Carnap's argument is perfectly cogent from a syntactic point of view. Conceived as a fragment of geometry concerned with finite sequences of abstract objects, the syntax of language simply has no resources to produce a philosophically interesting theory of truth. This is merely a special case of the more general deflationary effect that syntacticism has on metaphysics. Once the decision is made to treat language as mere syntax, no substantive metaphysical conclusions can be drawn on the basis of an analysis of language.[16]

16 The obvious caveat, and one that might bear considerable scrutiny, is the metaphysics of syntax itself. I will not pursue this theme in the present work.

Coffa misunderstands Carnap's argument as turning on an epistemic claim. We cannot know the truth-value of a sentence by knowing only its syntactic properties; therefore truth cannot be a syntactic property. This leads Coffa to identify verificationism as the root of Carnap's 'mistake' in thinking that a syntactic theory cannot provide a definition of truth. But verificationism is very clearly not at issue for Carnap. First, verification is not required for knowledge of analytic statements in general; the principle of verification only ever applies to synthetic sentences. Second, Carnap allows that there may be unverifiable analytic statements, e.g. there may be valid mathematical conjectures that we can never prove.

The argument Carnap gives against identifying truth with a syntactic property does not turn on any principle of verification. Rather, the argument consists in a proof that truth for language L cannot be defined within L[17] together with the observation that an adequate definition of truth in a metalanguage would constitute a shift to semantics, thus going beyond the boundary of syntacticism. Hence, Carnap explicitly considers

---

[17] The proof proceeds as follows. (Described in Sarkar 1992, the proof that follows was originally presented in its simplified form in Noonan (1990).) Carnap starts with three assumptions concerning a truth predicate T: (i) Every expression has exactly one of three properties T, F, and N. (ii) A sentence A has T iff A. (iii) Any expressions of types N(E), T(E), or F(E) do not have property N. From these three assumptions Carnap then locates a contradiction: Let E_1 be $T(E_2)$ and let E_2 be $F(E_1)$. E_1 and E_2 constitute a version of the paradox of the liar—E_2 says E_1 is lying, E_1 says E_2 is telling the truth.

Here is Carnap's formal derivation of a contradiction from his liar paradox sentences. (i) and (iii) imply $T(F(E_1))$ or $F(F(E_1))$. Assume $T(F(E_1))$. Then by (ii), we get $F(E_1)$ which together with (i) and (ii) implies not E_1, hence not $T(E_2)$. Since E_2 is defined as $F(E_1)$, assumption (iii) rules out $N(E_2)$, so by assumption (i) we must infer $F(E_2)$. Since E_2 is defined as $F(E_1)$, this gives us $F(F(E_1))$. Therefore, we infer $F(F(E_1))$ by disjunctive syllogism. $F(F(E_1))$ implies not $F(E_1)$ from (i) and (ii). Since E_1 is defined as $T(E_2)$, we have not $F(T(E_2))$. (i) and (iii) imply $T(T(E_2))$ or $F(T(E_2))$. By disjunctive syllogism, infer $T(T(E_2))$. By (ii) this gives us $T(E_2)$. Replacing E_2 with its definition, we have $T(F(E_1))$, and so by (ii) we have $F(E_1)$. By definition of E_1 we have $F(T(E_2))$, from which not $T(E_2)$ follows by (i) and (ii). We thereby arrive at $T(E_2)$ and not $T(E_2)$, a contradiction.

Coffa's idea that truth can be defined in a syntactic metalanguage. Carnap writes (in the sentences leading up to the passage quoted by Coffa),

> [I]t is possible to proceed without incurring any contradiction by employing the predicates 'true (in S_1)' and 'false (in S_1)' in a syntax of S_1 which is not formulated in S_1 itself but in another language S_2. S_2 can, for instance, be obtained from S_1 by the addition of those two predicates as new primitive symbols and the erection of suitable primitive sentences relating to them, in the following way: 1. Every sentence of S_1 is either true or false. 2. No sentence of S_1 is at the same time both true and false. 3. If, in S1, [sentence] G_2 is a consequence of [a class of sentences] R_1, and if all sentences of R_1 are true, then G_2 is likewise true. A theory of this kind formulated in the manner of a syntax *would nevertheless not be a genuine syntax*. (p. 216, emphasis added)

Carnap explicitly rejects Coffa's claim that a definition of truth can be constructed in a syntactic metalanguage for a reason that Coffa should accept, namely that such a definition of truth would not be genuinely syntactic but would necessarily invoke semantic notions. This was Coffa's point in his discussion of Tarski's model-theoretic notion of definition: in order to make the adequacy conditions for a definition of truth precise, one must define a class of truths that is the target of for the definition to capture, and doing this necessarily requires going beyond the boundaries of syntacticism.

The closing section of Coffa's essay recounts a story that Carnap used to tell his students about the first time Tarski explained to him his theory of truth. "They were at a coffee-house," Coffa writes, "and Carnap challenged Tarski to explain how truth was defined for an empirical sentence such as 'this table is black'. Tarski answered that 'this table is black' is true iff this table is black; and then—Carnap explained—'the scales fell from my eyes'."

This story tells, I believe, the full story of Carnap's semantic turn, but not in the way that Coffa thinks. Coffa speculates that Carnap "had been *so* close to Tarski's idea" that hearing this account revealed that his own "extraordinarily convoluted and artificial methods that made it impossible to understand exactly what was going on" should give way to Tarski's clear-headed picture of the situation. I contend that Carnap's methods were not "extraordinarily convoluted," nor did they prevent him from understanding what was going on. Rather, the formal similarity of Carnap's logical apparatus to Tarski's allowed Carnap to understand that Tarski's semantic approach was sufficiently exact and internally consistent such that it might be counted as an acceptable formal theory of truth. Since Carnap was always guided by the principle that what has been demonstrated in an exact and consistent way may be accepted, he had no reason not to accept Tarski's ideas. It was not that Tarski revealed a failing of Carnap's system; it was merely that Carnap saw in Tarski's work the possibility of a mathematically rigorous theory of semantic truth.

4.2 Creath on Carnap's LSL

Like Coffa, Creath (1990) criticizes Carnap for implicitly embracing semantics while giving a faulty reason for rejecting a semantic theory of truth. Quoting Carnap's claim that truth and falsehood are not proper syntactical properties, Creath replies: "What kind of reason is this?! Of course truth is not syntactical in this sense, but the question is why philosophy should be restricted to syntax in so narrow a way. Coffa interprets the argument as verificationism, but it is not even that. In technical terms this is just plain

'goofy'. It is as though 'and' is not a logical term on the ground that it is not a purely logical matter whether birds sing *and* Caesar marched. Or that 'two' is not properly definable in logic because it is not a purely logical matter whether there are two toads in Transylvania." Clearly, Carnap has gotten Creath very riled up, but is Carnap's argument really plain 'goofy'?

The first thing to note is that Creath concedes that truth is not syntactical in the sense that Carnap is concerned with. Carnap excludes the notion of truth from logical syntax because truth is a semantic property *sine qua non* and the project of syntacticism excludes semantics from consideration. After recognizing this point, it is obviously a mistake to interpret Carnap's argument as an argument for adopting syntacticism. Carnap's argument is meant to justify the exclusion of truth from the syntactic point of view, not to justify the adoption of the syntactic point of view.

So what is Carnap's argument for syntacticism? Carnap does not trouble himself to give a refutation of the semantic point of view that transcends his own syntactic framework. From within the framework of syntacticism, Carnap does offer many considerations against giving a semantic theory of truth. He frames his position in LSL as assumed without argument: "The view that, as soon as claims to scientific qualifications are made, all that remains of philosophy is the logic of science, cannot be established here and will not be assumed in what follows. In this part of the book [Part V, "Philosophy and Logical Syntax"] we propose to examine the character of the sentences of the logic of science, and to show that they are syntactical sentences. For anyone who

shares with us the anti-metaphysical standpoint it will thereby be shown that all philosophical problems which have any meaning belong to syntax."

As prime examples of philosophical problems that turn out to be pseudo-problems when stated in the material mode, Carnap cites "those logical sentences which assert something about the meaning, content, or sense of sentences or linguistic expressions of any domain. These are also pseudo-object-sentences." (p. 285) Carnap goes on to argue that sentences about meaning, designation, representation, about-ness, and content are pseudo-object statements. "Especially important here are those sentences which express a relation of designation, that is to say, those in which one of the following expressions occurs: 'treats of', 'speaks about', 'means', 'signifies', 'names', 'is a name for', 'designates', and the like." (p. 289) All such statements are to be translated into the formal mode of speech, so that their objects are syntactic objects.

What is the justification for citing sentences composed of semantic terms as pseudo-object-sentences and requiring them to be translated into the formal mode? The justification comes from empiricism. So long as semantic objects and relations are not in the domain of an empirical science, sentences concerning them can only be properly regarded as part of the logic of science, which can only be investigated syntactically. One might raise an anachronistic objection against Carnap to the effect that semantics *is* an empirical science, but his reason for excluding semantics from the domain of science in 1934 are clear enough.

Carnap's conception of semantics as metaphysics and his opposition to it on those grounds in LSL might most clearly be demonstrated by a pair of passages in which

Carnap contrasts his view of the logic of science with Schlick's interpretation of Wittgenstein's view. Carnap states his position thus:

> The sentences of the logic of science are formulated as syntactical sentences about the language of science; but no new domain in addition to that of science itself is thereby created. The sentences of syntax are in part sentences of arithmetic, and in part sentences of physics, and they are only called syntactical because they are concerned with linguistic constructions, or, more specifically, with their formal structure. Syntax, pure and descriptive, is nothing more than the mathematics and physics of language. (p. 284)

Carnap is adverting to the fact that the logic of science is partly *descriptive* syntax. It is an admixture of pure syntax—in its characterization of formal structure—and descriptive syntax in its coordinative identification of physical entities with syntactic entities. When Carnap says that "no new domain in addition to that of science itself is thereby created" he is partly reflecting Neurath's physicalism about the language of science. The structure that makes up the language of science is a physical structure, already a part of the domain of physical science. Investigating it does not require positing any new metaphysical entities such as truth or meaning or content.

Carnap contrasts this picture with a quote from Schlick talking about Wittgenstein:

> Schlick interprets Wittgenstein's picture as follows: philosophy "is that activity by which the meaning of propositions is established or discovered"; it is a question of "what the propositions actually *mean*. The content, soul, and spirit of science naturally consists in what is ultimately *meant* by its sentences; the philosophical activity of rendering significant is thus the alpha and omega of all scientific knowledge."

This is precisely the kind of metaphysical rhetoric that Carnap's syntacticism is intended to combat. Moreover, it is precisely the sort of rationalist monopoly on scientific

knowledge that Neurath's physicalist syntacticism is crafted to undermine. Carnap's rejection of semantic truth is rooted here, in his opposition to the notion that philosophy has a special role that ascends to greater heights than science while simultaneously bestowing meaning on scientific discourse through rational insight into semantic objects and properties. Carnap's rationale for rejecting a semantic theory of truth is clearly not given by the argument that Creath calls 'goofy'. Creath is badly mixed up about Carnap's motivations.

Creath goes further in his criticism of Carnap's rejection of semantic truth. He writes, "The argument against truth is so bad that it is plausible to assume that Carnap was antecedently prejudiced against the concept of truth...What is sad about the whole episode is not only that truth is in fact entirely compatible with the conventionalism and pragmatism of *Logical Syntax*. Rather, it is that the background prejudices against truth actually *fly in the face* of the central lessons of that book, namely, its epistemic conventionalism and its ontological neutrality."

Pace Creath, syntacticism appears to be necessary for Carnap's ontological neutrality. Creath himself emphasizes that Carnap is a non-cognitivist about metaphysics. Carnap defends a non-cognitivist position by showing that there is no translation of metaphysical statements into the formal mode. Metaphysical problems lose their significance when they are treated as syntactical, and there is no other way to treat them. However, this kind of non-cognitivism about metaphysics cannot be maintained once semantics is embraced. There is no longer any rationale for disavowing the *meaningfulness* of metaphysical stipulation. Hence, once Carnap endorses semantics, he

is compelled to say that ontology is relative to a linguistic framework—see Carnap's (1950) "Empiricism, Semantics, and Ontology." This constitutes a major departure from the metaphysical non-cognitivism derived from the syntacticism of LSL. As I will show in chapter 6, Carnap (1950) is certainly not "basically a reiteration of [Carnap's] conventionalism and pragmatism and the ontological non-cognitivism that follows from it," as Creath maintains.

Before dropping the subject of semantics and ontological neutrality (for now) let us briefly reflect on our earlier discussion of what makes arithmetical statements true. From the syntactic point of view, this question makes no sense. This allows Carnap to coherently oppose an appeal to some domain of mathematical reality that stands apart from language. The motivation for this opposition disappears once a semantic theory for mathematics is embraced; ontological neutrality is lost when one admits e.g. that there are numbers (or something else) that make the sentences of arithmetic true. More generally, when one adopts a semantic theory of truth for a domain of discourse, one takes on ontological commitments in that domain. Syntacticism is able to remain ontologically neutral where a semantic theory cannot be neutral.

What should we make of Creath's claim that Carnap's syntacticism flies in the face of his 'epistemic conventionalism'? The question suffers from a faulty presupposition that Carnap (1934) is an epistemic conventionalist. Carnap wasn't a conventionalist about epistemology until the liberalization of empiricism, at which point even empiricism itself becomes a convention (Sarkar 2013). But at the time of LSL,

Carnap had not yet adopted this view.[18] So there is no sense to be made of the idea that syntacticism 'flies in the face' of the epistemic conventionalism of LSL.

Like Coffa, Creath also alleges that Carnap already embraces several semantic notions in LSL despite his best efforts. Creath writes, "Not only is the logical consequence relation itself semantical, as we use the term, but so are truth tables, interpretation, and analyticity, all of which Carnap discusses." Later Creath says that Carnap's definition of analyticity is especially clearly semantic in nature. Unfortunately, Creath doesn't defend these assertions at any length—one gets the impression that his target audience (perhaps including Coffa) already agrees with him—so it is up to us to extrapolate his thoughts.

To start with the simplest case: there is nothing inherently semantic in the structure of a truth table. One often finds the "T"s and "F"s of the truth table replaced with "1"s and "0"s. Semanticists may regard the "T"s and the "F"s as meaning true and false, respectively, but this fact does not impose a semantic interpretation on Carnap's treatment of truth tables within a theory of logical syntax. Even if on Creath's view according to which truth is a real part of the metaphysical universe and the relationship between the truth of a complex sentence is determined by the truth of its parts in the way that the truth tables might be taken to indicate, it is still inappropriate to regard Carnap's discussion of truth tables as committing him to semantics. That would be analogous to regarding Newton's theory of physics as committed to quantum mechanics on the grounds that the universe is really quantum mechanical. A semanticist should know that

[18] His first endorsement of the position occurs in the second half of Carnap (1936).

if a theory T is about phenomenon X and X is really Y, it does not follow that T is about Y.[19] So even if the truth tables really are semantic in nature, the fact that Carnap talks about them does not entail that his theory is committed to semantics.

Carnap discusses the practice of interpreting one language by the use of another in an explicit attempt to refute the claim that a general theory of syntax must employ concepts of meaning. According to his position, "The interpretation of a language is a translation and therefore something which can be formally represented; the construction and examination of interpretations belong to formal syntax." (p.228) On the one hand, the translation itself is to be constructed as pure syntax. On the other hand, the applicability of a pure syntactic translation to a pair of physical languages is necessarily a matter of descriptive syntax. A descriptive syntactic theory of translation must typically take historically known habits of speech into account in order to be regarded as acceptable. Hence historical facts can influence the choice of syntax language used for the translation task, as well as the choice of (analytic) coordinating definitions adopted. (p. 228) These aspects of interpretation are in line with the principle of tolerance. But Carnap is clear that the notion of translation itself is completely given in pure syntax. He concludes his discussion of interpretation by writing, "we have established the fact that even the questions which refer to the interpretation of language, and which appear, therefore, to be the very opposite of formal, can be handled within the domain of formal syntax." There is

---

[19] I suppose some semanticists would accept this entailment, namely those who regard the context created by "is about" to be extensional. Even in that case, it is still mistaken to characterize Carnap's LSL as committed to semantic truth on the grounds that he discusses truth tables, since commitment must be understood in some way that rules out the claim that Newton was committed to quantum mechanics.

no reason to think that Carnap's discussion of interpretation implicitly supposes anything about semantics. A philosopher who thinks that translation is essentially semantical would clearly disagree with Carnap.

I assume that Creath's emphasis on the supposedly semantic nature of Carnap's definition of analyticity derives from his reading of Coffa (1987); as such, I have already addressed this concern and will not discuss it further.

Why does Creath think that Carnap's logical consequence relation is semantic? Creath mentions that Carnap's notion of consequence is transfinite. Recall that a sentence S is a consequence of a set of sentences R iff the negation of S is incompatible with R. Consequence is transfinite in that S may be a consequence of a set of sentences R, but not of any finite subset of R. This is only possible in higher-order logics for which compactness fails; it thereby undermines one possible desiderata for a syntactic theory of logic, namely that all logical consequences should be derivable from a finite set of axioms in a finite number of steps. Carnap recognized the need to reject this desideratum in light of Gödel's incompleteness results. But does the shift from a finitary to a transfinitary consequence relation constitute a shift from syntacticism to semantics? I see no reason to think it does. There is nothing semantic about an infinite set of strings, nor is there anything intrinsically semantic about the idea that a string could be incompatible with an infinite set of strings without being incompatible with a finite subset of them.[20]

4.3 Oberdan on Carnap's Syntacticism

[20] One could give a model-theoretic illustration of such an incompatibility, but model theory is not necessary for the definition of transfinite incompatibility.

Oberdan (1992) approaches the issue of Carnap's resistance to semantics from a perspective that is much closer to the present essay, according to which Carnap's syntacticism is the key explanans. However, most of what Oberdan says is not cogent.

First, Oberdan suggests that Carnap resisted a semantic theory of truth because such would obliterate his distinction between the formal mode and the material mode. Oberdan cites a passage in which Carnap writes:

> The difference between the quasi-syntactical sentences and the others is connected with syntactical concepts and the concept 'true'. If one were to take 'true' as a syntactical term, then every sentence whatsoever...would be quasi-syntactical.

However, the fact that a truth predicate would make every sentence quasi-syntactical does not in itself obliterate the distinction between the formal and the material mode. Quasi-syntactical sentences are ambiguous between the formal and material mode; they can be translated into either. While the notion of truth may trivialize the notion of a quasi-syntactic sentence, this does not seem like a very strong reason for Carnap to reject the notion of truth, especially when we reflect on the fact that the principle of tolerance allows for languages that are entirely composed of quasi-syntactic sentences.

Oberdan only offers this explanation as a preliminary consideration, however, and then moves on to what he takes the real issue for Carnap to be. He writes, "what Carnap's argument purports to show is that the only consistent approach to the definition of truth entails a hierarchy of languages. And much of the attraction of the method of syntax...was that it provided the means for metalinguistic discourse without the hierarchy...[so] certainly a more likely reason why Carnap rejected the method of defining truth in a

higher meta-language is because it would vitiate the argument which originally motivated his syntactical investigations, viz., that significant meta-linguistic discourse was possible within a single language." So, according to Oberdan, the real reason that Carnap rejected a semantic theory of truth was that he wanted to maintain that there was no hierarchy of languages.

This explanation is purely psychological. If Oberdan is right, then Carnap rejected a semantic theory of truth because he *wanted* to give a single-level account of language. But Oberdan also goes on to say, "Thus construed, Carnap's argument against truth follows from his syntacticism". So in addition to being motivated to give a single-level account, apparently syntacticism requires such.

Both of these thoughts are clearly wrong. Carnap's syntacticism is not opposed to allowing a hierarchy of languages, and neither was Carnap (psychologically speaking). For one example, recall what Carnap says about the formalization of mathematics: "everything mathematical can be formalized, but mathematics cannot be exhausted by one system; it requires an infinite series of ever richer languages." (p. 222) Clearly, Carnap is not opposed to accepting an infinite series of ever-richer languages, and nothing in his theory of logical syntax rules out such a thing.

Oberdan's confusion on this point enters with his lack of recognition that Carnap's statements in his "Metalogik" lectures concern *descriptive* syntax. Oberdan cites these lectures as evidence for the claim that Carnap was concerned to show that logical syntax can be expressed within a single language. Indeed, Carnap does support this notion, in part because it is required for his friend Neurath's project of giving a

language of unified science. What is required for Neurath's project is only that a theory of descriptive syntax can coordinate languages from different levels of the hierarchy with some physical medium, which is not a problem. But there is no reason to think that Neurath's desideratum enforced any constraint on Carnap's thinking about pure syntax. In that domain, Carnap had no concern over allowing a hierarchy of languages.

Oberdan presses on, writing that, "by the time Carnap wrote his argument against truth, his syntacticism was already obsolete, out-moded by two developments. The first is the conventionalism expressed in the Principle of Tolerance, and the second is his comprehension of Gödel's theorems." I will consider each of these two developments in turn and interrogate the claim that each renders Carnap's syntacticism obsolete.

The claim that the principle of tolerance renders syntacticism obsolete is dubious at the outset. The principle of tolerance allows us to adopt any set of axioms that can be formulated within logical syntax, but it does not license positing the existence of semantic properties in addition to adopting a syntax language.

Oberdan characterizes the principle of tolerance in the following way:

> The dogmatic defense of any particular form of language is expressly prohibited, because the choice among alternative languages does not involve matters of objective fact, or questions of right and wrong, but turns only on pragmatic factors. And these considerations never concern languages taken in isolation, apart from their use to express theories and beliefs. Rather, the choice among competing language forms depends only on the simplicity, economy of expression, and so forth, of the resulting systematic formulation of science.

It should be noted that Carnap's statement of the principle of tolerance says nothing about how the choice of language should be made. Carnap was explicit in saying that languages can be adopted for any purpose or for no reason at all. It should also be

noted that Carnap's motivation for the principle of tolerance had nothing to do with the thought that 'matters of objective fact' are not involved in the choice among alternative languages. It is an objective fact (as much as anything is) that there are a number of different and incompatible geometries; likewise, it is a matter of objective fact that there are a number of different theories of pure syntax. The 'choice' between them is not different than the 'choice' between Euclidian geometry and some non-Euclidian geometry—such a 'choice' is not based on any pragmatic factors, because a language of pure syntax is not intrinsically 'for' anything.

How does the principle of tolerance undermine syntacticism? Oberdan writes:

> Theology entered syntax when the difference between 'definite' and 'indefinite' was extended to syntactical concepts, especially the concepts used to characterize the kinds of inferences permitted in a language. When the rules of a language permit only inferences in which a finite number of premises occur, the result is a *derivation* of the conclusion. If, however, the transformation rules allow inference from finite series of (not necessarily finite) classes of sentences, the conclusions drawn are *consequences* of their premises, and chains of inference of this sort are called '*consequence-series'*.

So the problem is, supposedly, that the principle of tolerance allows us to adopt indefinite languages, and indefinite languages are beyond the bounds of pure syntax in virtue of allowing transfinite inference. There is no argument provided for the claim that logical syntax gives way to semantics once transfinite inference is allowed, apart from the *italics* imposed on the words "derivation", "consequences", and "consequence-series". Oberdan apparently just assumes that languages that go beyond the constructivist limitations of Language I are no longer fully syntactic. But as I have indicated, there is no reason to think that allowing for transfinitary inference entails a semantic framework.

Moving on from this 'problem', Oberdan turns to Carnap's definition of analyticity. Oberdan centers his investigation on Carnap's notion of valuation. He identifies the difference between this notion and Tarski's notion of satisfaction. He considers an example sentence "5 is a prime number" and notes that the difference between Carnap's approach and Tarski's consists in the fact that Carnap's valuation is concerned only with expression types—for example, the numeral "5"—while Tarski's is concerned with extra-linguistic entities—for example, the number 5. Having clearly articulated the distinction, Oberdan makes the following curious argument:

> [T]he only salient difference between Tarski's semantical approach and Carnap's syntactical method is the distinction between an entity, a number in the example, and its name, or a numeral in the present case. But imagine what would happen if Carnap, like Tarski, had applied the semantical satisfaction relation rather than his syntactical valuation to formulae of the object-language. The sentences expressing the application would belong neither to the formal mode – for they would contain expressions referring to extra-linguistic entities – nor to the real-object mode – since they contain expressions referring to (possibly other) linguistic expressions. So the preservation of syntacticism appears to rest on the rather slender basis of the distinction between names and the things they name."

Oberdan then moves on as if the point has been made. But there is no case here against syntacticism. Clearly, the distinction between names and the things they name is not a slender basis on which to rest the preservation of syntacticism! Virtually no philosopher has ever maintained that there is no distinction between a thing and the name of that thing. Certainly a semantic theory requires such a distinction (perhaps even more than syntacticism does). Oberdan has not argued for a collapse of this distinction. The syntactic view advocated by Carnap may be parallel in some sense to the semantic view advocated by Tarski, but the shortest transversal between the two views is quite long.

Oberdan's observation does nothing to make the case that Carnap's syntacticism is threatened by its superficial similarity to Tarski's semantic theory of truth.

The last objection Oberdan raises against Carnap's syntacticism concerns the indefinability of "analytic" within Language II. Oberdan points out that Carnap is concerned to show that a hierarchy of languages can be conceived as concentric parts of a single language. Oberdan then bizarrely claims that "construing the hierarchy of languages intra-linguistically allowed Carnap to preserve, for the last time, the syntacticist theme that dominated his thought in the early thirties." He argues that, since analyticity cannot be defined within Language II but implicitly requires an infinite hierarchy of meta-languages, Carnap's syntacticism collapses.

I must admit that I do not understand this argument. Why does syntacticism require denying the hierarchy of languages? What danger would syntacticism be saved from if it were possible to construe the hierarchy of languages as concentric regions of a single language? The fact that "analytic" is not definable within Language II is not a problem for syntacticism. Moreover, Carnap accepts that there is an infinite hierarchy of languages, each of which is a purely syntactic object. It is true that there is no language such that "analyticity" can be defined in that language, but for every instance of an analytic sentence S, there is a language in which "analyticity" can be defined syntactically such that S is analytic. Carnap explicitly endorsed this and recognized it as consistent with his syntacticism.

Carnap's desire to show that the hierarchy of languages can be encoded into 'concentric' regions of a single language is related to Neurath's project of developing a

unified science. If there is only one physical medium for the formulation of the language of science, and that language is able to talk about itself in every instance, then it must be possible for every part of the language to be the object of some physical discourse. If the indefinability of 'analyticity' makes any trouble, it is trouble for Neurath's conception of the physical language. Perhaps there just is not enough physical substance to encode an infinite hierarchy of languages.

Carnap would certainly grant the empirical possibility that the physical language medium turns out not to allow for an infinite number of expressions—this is a question about the physical world. If that were the case, it would be impossible to coordinate all the elements of all of the infinite hierarchy of syntactic languages with elements of the physical system. But the practical need to do such coordination is non-existent. (Neurath would also not fret about the fact that the infinite number of mathematical languages cannot be identified with some set of physical phenomena.) In any case, none of these concerns carry any significant weight for Carnap, who identifies pure syntax with a fragment of arithmetic and is thus relatively unconcerned about the limitations of physical language mediums.

4.4 Carnap's reason for abandoning syntacticism

When Carnap (1963) reflects on his semantic turn, he gives three reasons for his conversion. First, he says that it was brought on through conversations with Tarski and Gödel in which he realized that a precise formulation of semantic concepts was possible. Second, he says he realized that since it was okay to talk about facts and also to talk

about language it was okay to talk about the relations between facts and language. His third reason is pragmatic: "To me the usefulness of semantics for philosophy was so obvious that I believed no further arguments were required and it was sufficient to list a great number of customary concepts of a semantical nature[.]"

The only sense in which the seeds of semantics were present in LSL is the sense in which the logical investigations of that book prepared Carnap, given his understanding of logic, to grasp and accept Tarski's formulation of a semantic theory of truth. Carnap does *not* say: 'once I understood Tarski's theory, I realized that my own work was already committed to semantics.' Carnap remains clear that LSL was concerned with syntax only. Moreover, none of the three reasons that Carnap offers for his conversion indicates that Carnap's theory of logical syntax was flawed. At most, the Carnap of the semantic turn would say that LSL was incomplete in failing to encompass the semantic nature of language in addition to its syntactic nature.

In accepting that language had a semantic nature, Carnap was parting company with Neurath. He failed to be moved by Neurath's protestations. To those protestations, we now turn.

Chapter 5: Neurath's Opposition to Semantic Truth

Neurath's objections to the semantic conception of truth were considered obscure by his opponents, including Carnap, and remain obscure to interpreters today. This chapter attempts to clarify Neurath's opposition to Tarski's theory of truth and Carnap's acceptance of it. I will address three competing views about how to understand Neurath's opposition, given by Mormann (1999), Uebel (2001, 2007, 2009), and Mancosu (2008).

5.1 Mormann on Neurath

Mormann (1999) attempts to explain Neurath's opposition to a semantic theory of truth as an opposition to the over-application of mathematics and formal logic in the theory of empirical knowledge. Mormann conceives Neurath's grievance with this over-application as stemming from his anti-Cartesianism. I agree that Neurath's opposition to semantics is related to his opposition to an overemphasis on rationalist insight, but I do not think that Mormann is right to think that Neurath was opposed to formal semantics qua formal theory.

Mormann presents two aspects of Neurath's philosophy as components of the explanation for his rejection of formal semantics. The first is that the language of unified science cannot be made formally precise due to the historical necessity of using imprecise terms from ordinary common language in doing science. Mormann regards this as Neurath's "main and most radical attack" on Cartesianism, concerned with the Cartesian doctrine of (what Mormann calls) *transparentism*, which states: "The system of (scientific and philosophical) knowledge can be expressed in a language whose concepts

refer to clear and distinct ideas." Neurath rejects the idea that the expressions of science have stable meanings that are transparent to us. All concepts are open to revision because there is no final state of completed science. Mormann writes, "The language of science always remains a provisional language dependent on unclear and 'suspicious' concepts, [Neurath's] notorious 'Ballungen'." So the lack of transparency of the language of unified science is to be derived from the inexorable need to invoke Ballungen in scientific discourse.

The second aspect of Neurath's philosophy that contributes to his rejection of formal semantics, according to Mormann, is Neurath's adherence to a philosophical tradition that regards language as a universal medium. This is half of a distinction discussed by Hintikka (1996). The other half of the distinction is a tradition that treats languages as calculi, such that different languages can be formulated, invented, tinkered with. When language is treated as a universal medium, it makes no sense to talk of constructing other languages. Mormann writes, "According to [Neurath], there is only one language – the physicalist language." Proponents of the language-as-universal-medium tradition—Frege and Wittgenstein among them—typically maintain that speaking about language is not possible, but (as Mormann notes) Neurath rejects this thesis. As we saw earlier, Neurath maintains that the physical language can 'speak about' itself without contradiction. Nevertheless, according to Mormann, Neurath is to be considered a proponent of the language-as-universal-medium position, and this is supposed to help explain his rejection of formal semantics.

It is worth noting that Mormann, like Carnap in his intellectual autobiography, mischaracterizes Neurath's physicalism. Mormann says, "According to Neurath, physicalism, put roughly, is the thesis that all intellectually respectable concepts can be defined ultimately and entirely in terms of physicalist concepts and/or the concepts of logic and mathematics." There is no defense given for this interpretation; perhaps Mormann regards it as the received view of Neurath's physicalism. As I have argued in chapter 2, Neurath's physicalism should be understood as a two-part thesis: first, that language itself is a physical structure, and second, as a socio-political maxim that the language of unified science should be constructed in such a way that it only incorporates expressions that are either controlled directly by the senses or feature in inferences from observation statements to prediction statements. There is no requirement of definability.

It seems that Mormann regards Neurath's Ballungen and his commitment to the language-as-universal-medium hypothesis as independent reasons for the rejection of semantics. In the first case, formal semantics is inapplicable to the language of universal science because that language contains Ballungen. Since formal semantics is precise, and Ballungen are necessarily imprecise, formal semantics cannot be applied to Ballungen; hence, it cannot be applied to unified science. In the second case, the language-as-universal-medium hypothesis rules out the possibility of meta-languages, which are essential for a Tarskian approach to semantics.

Neither of these explanations can be quite right. First, Tarski's formal semantics is not more precise than a language of science that contains Ballungen. A typical example of Ballungen at work in unified science for Neurath would be a protocol sentence of the

form: “A clock is seen by Otto with both hands of the clock pointing to the numeral 12.” This protocol sentence contains Ballungen, such as “Otto”, “clock”, “hands of the clock”, “pointing”, “seen”, and so on. Now, consider Tarski’s semantic treatment of this protocol sentence: “A clock is seen by Otto with both hands of the clock pointing to the numeral 12” is true iff a clock is seen by Otto with both hands of the clock pointing to the numeral ‘12’. Note that Tarski’s definition of truth invokes the same expressions—including the Ballungen—that the sentence itself uses. The semantic interpretation invokes the very same Ballungen as the object sentence, and so it is exactly as indeterminate. There is no mismatch in the precision of Tarski’s semantics with the precision of the language of unified science.[21]

Mormann’s other explanation seems to be faulty as well, because Neurath explicitly rejects the language-as-universal-medium hypothesis. He writes,

> It has rightly been pointed out that it would be better to speak of the languages of physicalism, since more than one language could satisfy the conditions demanded by physicalism in principle. Here again it is shown that logical empiricism takes pains to replace absolute formulations with their ‘either-or’ by several possible ‘suggestions’. In practice, however, we advocates of this ‘tolerance’ will make some pronouncements with greater definiteness than our own principles allow: then we just have to ‘relativise’ each other. (Neurath 1936e/1985 p. 165)

This position is in keeping with Neurath’s Duhemian commitments, as well as the theoretical commitments of physicalism as I have presented it. Neurath had no reason to

---

[21] I granted earlier that Neurath’s misgivings about incorporating the term “true” into the language of unified science was partly to be explained by his awareness of Ballungen. That explanation, it seems, cannot be directly applied to answer the (different) question of why Neurath was opposed to Tarski’s formal theory of semantics.

regard the physical language of unified science as the only language; nor did he need to maintain that there is only one possible physicalist language.

In addition to this pluralist position about the language of physicalism, Neurath also accepts that the formal sciences can develop calculi of the sort that Tarski and Carnap are interested in. Indeed, Neurath maintains that investigations into formal languages such as those carried out by Russell, Whitehead, and Carnap are necessary prerequisites for developing a unified science. He encourages Carnap to develop a logical syntax of language that could serve as the basis for a science of the physicalist language.

Confusingly, Mormann recognizes this point. He remarks that Neurath was amenable to Tarski's formal semantics as long as it was confined to the domain of formal languages. That Neurath did accept Tarski's definition of truth, so long as it is confined to formal mathematical discourse, shows that he accepted something like that language-as-calculi view for formal languages. That he did accept some form of this position is apparent in the correspondence between Neurath and Tarski (presented in Mancosu 2008; see below).

Neurath's opposition to semantics does not stem from his commitment to Ballungen, nor is he opposed on the grounds that the physicalist language is a universal medium and therefore meta-languages are impossible. We must look elsewhere for the correct explanation.

5.2 Uebel on Filling in Neurath's 'blind-spot'

Uebel (2007) regards Neurath's opposition to Carnap's semantic turn as a philosophical blunder that could have been avoided. He writes that "Neurath's blind-spot was...to think that one could do without the concept of truth, and thus without a theory of truth—however minimal—altogether." However, rather than regarding this 'blind-spot' as theoretically motivated by Neurath's conception of science, Uebel "feels pressed to field reasons like differences in temperament [between Neurath and Carnap] to explain Neurath's resolute resistance to truth-talk". Treating Neurath's resistance to semantics as the result of temperament rather than as an important and inextricable consequence of his physicalism would, according to Uebel, open the possibility of keeping Neurath "at the table of the logical empiricist project" by allowing us to charitably amend Neurath's view by adding a semantic component.

Uebel proposes that the right way to save Neurath from philosophical irrelevance is to supplement his physicalism with a disquotational theory of truth. This, Uebel says, would preserve the spirit of Neurath's philosophy while saving him from his erroneous rejection of semantics. Neurath's primary motivation for rejecting semantics, as Uebel identifies, is the fear that semantics will lead to a resurgence of metaphysics. (Uebel recognizes this as Neurath's primary motivation, yet somehow also maintains that his rejection of semantics is only a matter of temperament. I think Neurath's fear is not merely a matter of temperament but instead follows directly from Neurath's understanding of metaphysics and its connection with semantics.) Amending Neurath's

physicalism with a disquotational theory of truth is thought to be 'in the spirit' of Neurath's philosophy insofar as it "retain[s] his anti-metaphysical animus."[22]

I do not think Neurath's physicalism can be amended with a disquotational theory of truth without undermining his anti-metaphysical animus. A disquotational theory of truth, while anti-metaphysical in some respects, still leads to metaphysics in just the way Neurath was alert to. This fact shows that Neurath's rejection of semantics was not a matter of temperament, nor was it an unfortunate accident. Rather, it was a necessary cornerstone of his physicalism and his anti-metaphysical stance.

A disquotational theory of truth purports to eliminate one metaphysical domain that Neurath was opposed to, namely the domain of semantic relations. The essential difference between a disquotational theory of truth, or deflationary theories of truth more generally, and a non-deflationary or robust theory of truth is that the latter but not the former claims that there is an informative theory about the nature of truth (Field 1994, David 1994, Horwich 1998, Hill 2002). Deflationary theories of semantics, including disquotational theories, eliminate a certain class of metaphysical explananda. The eliminated explananda include questions such as: How does a sentence get its truth conditions? What is the nature of reference? Is truth identical with some physical property? Is truth correspondence with reality? Does every true sentence correspond to a

---

[22] Uebel passes quickly over the claim that Neurath should be happy to accept a disquotational theory of truth and states that the only technical obstacle to adding such a theory to Neurath's philosophy consists in addressing the issues raised by Mormann (1999). I don't think the issues raised by Mormann constitute any technical obstacle to adding disquotationalism to Neurath's physicalism. The obstacle is not technical but philosophical: to add a theory of truth, even a disquotational one, is to reject Neurath's physicalism and undermine his anti-metaphysical animus.

fact? Neurath would surely regard these questions as pseudo-scientific and metaphysical. Insofar as a disquotational theory of truth undermines these questions, it is amenable to Neurath.

However, Neurath's syntacticism functioned to undermine more than just the metaphysics of truth, reference, fact and reality. Neurath's syntacticism provided a vaccination against a whole range of metaphysical claims. For one instance, consider the disquotational yet semantic claim that "Water is identical with H_2O" is true iff water is identical with H_2O. The left hand of the biconditional is concerned with language, but the right hand side is not. Once this biconditional is part of unified science, the way is open to questions of scientific realism and scientific essentialism. Is there really such a thing as a hydrogen atom? Is the essence of water its chemical composition? These metaphysical issues and many more are raised by even a disquotational theory of truth. (We will return to the variety of metaphysical questions that are opened by semantics in chapter 6.)

Note that disquotationalism is significantly different from the non-semantic redefinition of "truth" that is already part of Neurath's conception of physicalism. Neurath allows that "true" may be re-conceived as applicable (in our global scientific community) to a sentence just incase the sentence is part of our encyclopedia. Thus, the program of unified science admits of the following rule: if "'S' is true" is part of the encyclopedia, add "S" to the encyclopedia. It must be emphasized that a rule of this kind is **not** what is given by a disquotational theory. A disquotational theory provides equivalences between statements about language and statements that are not about language. The difference between Neurath's syntactic re-definition of "truth" and a

disquotational theory of truth is as significant as the difference between Carnap's syntactic theory of valuation and Tarski's semantic theory of satisfaction.

Disquotationalism raises another metaphysical issue for Neurath's physicalism. A semantic theory of truth, even a disquotational one, countermands a primary aspect of Neurath's physicalism, viz. the claim that expressions of a language are physical structures with no special properties—no properties that are not had by any other physical structure. In order to fit a cogent disquotational theory into Neurath's physicalism, that theory must impose a restriction on the substitution class of physical objects. A substitution class is the set of things that can be substituted for the variable S in the T-schema: "S" is true iff S. The only things that can be substituted are *meaningful* sentences of some target language. It doesn't make any sense to substitute e.g. a chair for S. Thus, disquotationalism entails that linguistic expressions have properties—semantic properties concerning their meaningfulness—that no other physical systems have.

These considerations refute Uebel's claim that Neurath's resistance to semantics was merely temperamental. Neurath's syntacticism is the heart of his physicalism. To reject it in favor of semantics would be to give an entirely different philosophical picture, one that is much less radical and relatively incapable of undermining metaphysics to the degree that Neurath intended.

Does this mean that Neurath has been excused from the table of serious logical empiricists? I do not believe it does. Semantic realism is certainly very integrated into philosophy. It is the ubiquitous legacy of the linguistic turn, infiltrating almost all of contemporary philosophical discourse. It is not a coincidence that the ascendency of

semantics coincides with the resurgence of metaphysics, as Neurath forewarned. Uebel notes that Neurath was right to predict that metaphysics would be on the rise if Carnap and Tarski had their way. Therefore, it is not unreasonable that serious logical empiricists weary of the present era of metaphysics should reconsider Neurath's proposal of physicalist syntacticism.

5.3 Mancosu on Neurath, Carnap and Tarski

Mancosu (2008) provides a very careful investigation of Neurath's expression of frustration with and opposition to the semantic conception of truth through a detailed presentation of Neurath's work together with correspondences between Neurath and Tarski, Carnap, and Kokoszynska. Mancosu concludes that Neurath's opposition was based on the considerations that led him to propose re-defining "truth" as a label for indicating that a sentence was part of the encyclopedia, as well as arguments against the semantic definition of truth that were derived from those considerations. The heart of the issue is that semantic truth is a threat to empiricism.

The picture developed in the present essay is largely in agreement with Mancosu's assessment. All that is missing from Mancosu's explanation is what is missing from Neurath's presentation of his own concerns: an explicit articulation of how semantics leads to metaphysics.

It was this point that Carnap could not grasp. According to Mancosu's recounting of Carnap's final official exchange with Neurath on the topic, Carnap (in a typescript dated 7/18/1937; cf. Mancosu 2008) presents four theses:

1. The semantical conception of truth is correct and unobjectionable;
2. It cannot be replaced by merely syntactical method;
3. It is useful and important;
4. It is in agreement with the concept of truth used in ordinary language.

Mancosu does not speculate on which of these theses Neurath would take issue with. It seems that Neurath would not in fact take issue with the first thesis, so long as the definition of truth is constrained so that it is only applied to formal languages. A letter to Tarski from Neurath (dated March 24, 1936) is revealing about his thoughts on the matter. In this letter Neurath writes,

> The restrictions you impose on the concept of truth will not be observed and your formulations will be used for all kinds of metaphysical speculations. (Quoted in Mancosu 2008, p. 197)

At another place in the letter Neurath writes,

> ...insofar as your terminological choice suggests objectionable consequences, it has perhaps not come about independently of these consequences. On the one hand one emphasizes that this concept of truth holds only for formalized languages. On the other hand the concept of truth is of practical interest precisely in non formalized domains. For this reason, if one is not simply to get rid of the term, I am in favor of my terminology, for the latter remains applicable also in non formalized domains. By contrast the terminology you and Lutman[-Kokoszynska] use leads to bad things when it is applied to non formalized domains. (quoted in Mancosu 2008, p. 198)

There are three important insights to be drawn from this passage. First, Neurath accepts that Tarski's definition of truth holds for formalized languages. This is perfectly cogent for Neurath, who is content to allow mathematics and formal language to feature in the portion of the program of unified science that connects observation statements with predictions. There is no in-principle restriction on the type of expression or methods that

may be employed in the investigation of the formal sciences; this is because the sentences that intervene between observation and prediction are not statements but directives for transforming sentences into sentences. Hence, Neurath is not worried about Tarski's concept of truth as a part of mathematical formalism.

Second, Neurath denies that "truth" as defined by Tarski can be applied to non-formalized domains. This denial must be understood in the context of Neurath's physicalism. Since language is only a physical structure, characterized by physical syntax, the use of a term in a language cannot be excluded on purely rational grounds. Neurath's rejection of Tarski's definition of "truth" can only be pragmatic by Neurath's own lights, part of the social engineering project. It is a confusion to think of Neurath's prescribed restriction of the use of the term "truth" as a denial that "truth" applies (in some semantic sense) to the domain of 'facts' as opposed to the domain of logical discourse. The exclusion is a pragmatic recommendation about how we should use our language. Yet this pragmatic recommendation is grounded in Neurath's politico-epistemological project. Neurath is convinced that using the word "truth" according to Tarski's semantic definition will put that word out of the control of the senses, and do so in a way that does not contribute to making predictions but will lead instead to the production of metaphysical speculation and debate and a reinforcement of the social value of elite rational insight. Thus, Tarski's truth definition is dangerous to empiricism in that if it is accepted by the masses, the program of unified science will be rendered dysfunctional.

The third insight to be drawn from Neurath's comments to Tarski is that, while Neurath recognizes that Tarski has officially restricted his definition to formal domains, he also suspects Tarski of having an ulterior motive in calling his formal concept "truth" rather than something less likely to be applied outside of the formal domain. The notion of truth is "of practical interest precisely in non-formal domains," so someone naming a strictly formal concept "truth" is deliberately trying to mislead people. This is what Neurath is about when he says that Tarski's terminological choice has not come about independently of the objectionable consequences that his choice will have in the public domain. Neurath is explicitly identifying Tarski as a rationalist public enemy acting with political intent. Neurath thinks Tarski's choice of the term "truth" for his formal concept is intentionally designed to to induce people to apply the his definition in non-formal domains.

So, while Neurath should be inclined to accept Carnap's first thesis that the semantic conception of truth is 'correct and unobjectionable' on the grounds that it is correct and unobjectionable within formal domains, and so technically admissible into the program of unified science, he would be inclined to reject it on pragmatic politico-epistemological grounds. You can't give the people the word "truth" with a formal definition and expect them to use it responsibly.

Carnap's second thesis—that the semantic notion of truth cannot be replaced by syntactic methods—is something that Neurath can also agree with in a sense. Both Carnap and Neurath recognize that truth is not a syntactic concept. Neurath's recommendation to re-define "truth" to mean "included in or entailed by sentences

included in the encyclopedia" is not intended as a replacement in the sense of an accurate explication of the original concept. On the other hand, Neurath would have disagreed with Carnap's second thesis from a social engineering perspective. There is no in principle reason why we could not, as a society, decide to replace the common practice of using the term "true" with semantic connotations with a new practice that conforms to Neurath's recommended change of usage reflecting a fully syntactic definition.

Carnap's fourth thesis—that the semantic conception of truth conforms to common usage—is something that Neurath argued against. In his correspondence with Kokoszynska (as presented in Mancosu 2008), Neurath writes: "The Tarski-Lutman definition does not correspond to the ordinary usage in any privileged way (historical question." Neurath concerns himself with arguing that there are many patterns of usage of the term "true" that could be identified as a matter of historical fact. Neurath is not convinced that the semantic definition of truth is in agreement with the commonsense concept.

Yet, from the social engineering perspective, historical questions of how the term has been used in the past are not decisively relevant to the question of how we should use the term now and in the future. It is thus not directly relevant to Neurath's position whether Carnap's fourth thesis is accepted. Perhaps common usage is in agreement with Tarski's definition; it is another question entirely whether, on reflection, we should decide to continue use the term in the way that it has been used historically or adopt a new procedure.

So, Neurath and Carnap do not disagree in any deep way about theses one, two, or four. The substantive disagreement between Carnap and Neurath can therefore be located entirely in thesis three: that the semantic conception of truth is useful and important. This is a pragmatic issue. To what use can the semantic concept of truth be put? Carnap is interested in analyzing philosophical notions. By his own lights, these are notions that were taken off the table by syntacticism. Neurath holds that the philosophical analyses that are made possible by the semantic conception of truth belong to the realm of metaphysics. It is not important to carry out such analyses—in fact, the opposite is true: it is important, for the sake of placing epistemic power in the hands of the people, that we resist placing any importance on semantic analyses.

It is in this light that Neurath sees Carnap's semantic turn. He believes that Carnap has been led astray by Tarski's theory of truth and been seduced into seeing philosophical importance in metaphysical questions to the detriment of empiricism and, more broadly, to the detriment of the general population struggling for social equality. Hence it is in all sincerity that Neurath writes to Carnap in defeat (dated 1/15/1943, as quoted in Mancosu 2008):

> I am really depressed to see here all the Aristotelian metaphysics in full glint and glamour, bewitching my dear friend Carnap through and through. As often, a formalistic drapery and hangings seduce logically-minded people, as you are very much...It is really stimulating to see how the Roman Catholic Scholasticism finds its way into our logical studies, which have been devoted to empiricism.
>
> The Scholasticism created Brentanoism, Brentano begot Twardowski, Twardowski begot Kotarbinski, Lukasiewicz (you know his direct relations to the Neo-Scholasticism in Poland), both together begot now Tarski etc., and now they are God fathers of OUR Carnap too; in this way Thomas Aquinas enters from another door Chicago...

Chapter 6: How Neurath's Nightmare Has Come True

Neurath warned that engaging in semantic analyses of language would lead to a new era of metaphysics. In this chapter I argue that Neurath was right. Many of the most influential metaphysical ideas of the 20th century were articulated and defended through reasoning about semantics. What follows are a number of case studies in metaphysics, with special attention paid to the role that semantics plays in establishing metaphysical conclusions. In each case, I sketch the situation from the point of view of Neurath's physicalist syntacticism and indicate some anti-metaphysical conclusions that might have been reached if not for the semantic turn.

6.1 Empiricism Semantics and Ontology

Carnap's (1950) stated purpose in this quintessential essay of his semantical period is to show that ontological disputes are pseudo-questions. His strategy is to demonstrate this by showing that questions of ontology are trivial when raised inside a linguistic framework and cannot be coherently raised at all when no linguistic framework is assumed. But from the perspective of Neurathian physicalist syntacticism, Carnap has already taken the plunge into deep metaphysical waters by treating linguistic frameworks as semantically interpreted.

A linguistic framework for a set of entities is constructed by introducing a system of ways of speaking, subject to rules. This procedure is not very different from the construction of a logical syntax, except for the crucial fact that a linguistic framework includes a semantic interpretation. Thus, when one introduces a linguistic framework, the

expressions of the framework are assigned designations: sentences designate propositions, predicates designate properties, and names designate things.

Carnap's discussion provides an explicit characterization of the link from semantics to ontology. He walks us through an example. Start by accepting the linguistic framework of numbers. Within this framework, the statement "Five is a number" is analytic. From the fact that "Five is a number" is analytic, it follows that "There is some n such that n is a number" is also analytic.

Neurath would be perfectly happy to countenance rules of transformation of this kind if they were to be construed syntactically. From the perspective of Neurath's physicalism, we very well might include the string "There is some n such that n is a number" in the encyclopedia by inferring it from "Five is a number". But this statement, construed as part of the program of unified science, does not express any more than its own physical syntax. In syntax there is no ontology, no domain of facts beyond the language itself.

Carnap, however, has taken his semantic turn. Now, analytic sentences are *true*. So from the fact that "There is some n such that n is a number" is analytic, it follows that it is *true* that there are numbers. More generally, accepting a framework means regarding the analytic sentences of that framework as true. It is not surprising, then, that questions of existence can be settled by adopting frameworks, for one can adopt a framework that asserts the truth or falsity of the answers to those questions.

Carnap presents this picture as if it were deflationary. He says, "Nobody who meant the question 'are there numbers?' in the internal sense would either assert or even

seriously consider a negative answer." But that is because by definition, in accepting the framework of numbers, they have accepted that "Numbers exist" is true and thus accepted that there are numbers. From a perspective that already accepts semantics, Carnap's remarks can appear to be deflationary, but from the perspective of physicalist syntacticism it is clear that adopting a linguistic framework—thereby accepting the truth of sentences that assert the existence of extra-linguistic entities—is to stray very far from the strict empiricism of LSL.

The contrast between internal questions and external questions is rendered equally trivial from the perspective of syntacticism. Carnap observes that once we accept the number framework, and thereby accept that there are numbers, the question of whether there are *really* numbers no longer makes sense. Carnap thinks that this conclusion indicates that questions of reality are pseudo-questions. But this is not how the syntacticist sees the situation. From Neurath's point of view, once we have accepted Carnap's semantic framework and endorsed the idea that there are true statements of e.g. arithmetic, we have already taken the plunge. In answering the question: do I accept the framework of numbers? I have committed myself to answers to all of the relevant metaphysical questions about numbers. The person who accepts the framework of numbers thinks it is true that numbers exist; the person who does not accept that framework (or another that entails the same truths) will not think that numbers exist. The fact that we cannot identify a meaningful question of the reality of numbers that is separate from the question of whether to adopt a linguistic framework does not show that metaphysics has been avoided. All that it shows is that answering the question of which

framework to adopt fully determines our answers to metaphysical questions. This is because accepting linguistic frameworks requires accepting truths about non-linguistic reality. Despite his intentions to the contrary, Carnap demonstrates the bridge from semantics to metaphysics.

Let us consider a likely objection to this interpretation. A defender of Carnap might say: Ontology is relative to a linguistic framework, and a linguistic framework is merely a set of rules for speaking, so how can adopting a linguistic framework carry any metaphysical significance?

We reply, with Neurath: If linguistic frameworks were characterized syntactically—if they were really merely rules for the manipulation of physical structures—then frameworks would not carry any metaphysical significance. However, we have passed into the semantic turn, and now Carnap's linguistic frameworks have semantic properties. When the rules you adopt entail that a sentence analytic, that sentence is now to be regarded as true. So adopting a framework is not just a matter of accepting rules for speaking; it is also a commitment to truths that go beyond language.

A defender of Carnap might persist, arguing that Carnap's logical truth and its attendant ontology is not a significant commitment to metaphysics. How could it be significant, given that *anything* can be regarded as a logical truth? Carnap's position in "Empiricism Semantics and Ontology" is the result of combining the principle of tolerance and its attendant conception of analyticity with a semantic theory of truth. The result is the view that any sentence that can be assumed as an axiom can be regarded as

true. Thus, any sentence can be regarded as logically true. Doesn't this render the notion of logical truth harmless by making it trivial?

From Neurath's perspective, the question of the harm involved is not to be assessed in terms of whether the notion of truth presupposed is trivial in the sense that it may be adopted for any sentence, but whether accepting this account of logical truth gives rise to metaphysical speculation, debate, and conviction. In these respects, Carnap's view constitutes harmful metaphysics. Consider some of the questions that it raises. If you and I adopt different frameworks, then what is logically true for me is not logically true for you; is logical truth therefore relative to a person at a time? I can adopt a framework according to which it is logically true that there are no electrons; is the existence of electrons then to be decided by my choice of words? Carnap mentions that we might refrain from speaking and thereby reject all linguistic frameworks; does this mean that the existence of numbers requires speech? Even metaphysical questions that Carnap explicitly aims to rule out are still open for discussion: Are numbers real, if their existence depends on the adoption of a linguistic framework?

It is clear that accepting a theory of semantic truth led Carnap into a realm that was not available from the perspective of LSL. But even in his semantic phase Carnap was still committed to empiricism. He worked diligently to contain the metaphysical implications of the semantic framework he endorsed. Philosophers more inclined to trust in rational insight were less reserved about using the semantic framework. It was only a matter of time until rational insights about semantics would drive philosophical discourse much farther into the depths of metaphysics.

6.2 Rigid Designators, Necessary Identity and Essence

The first line of Kripke's (1980) *Naming and Necessity* is: "I hope that some people see the connection between the two topics in the title." The connection is made explicit through Kripke's extremely influential semantic arguments concerning the content of names as designating the same individual across all possible worlds and its relation to knowledge and metaphysical necessity.

Kripke (1980) begins with a flurry of considerations about the semantic content of proper names and descriptions. The content of a name cannot be identified with a description or collection of descriptions that pick out the individual named. A description may be satisfied by different individuals in different possible worlds, but a name picks out the same individual in every possible world.

Consequently, a correct semantic theory will never entail that some description of S is necessarily true in virtue of the fact that S is the semantic value of the name "N". Yet it is possible to give descriptions of an individual that are necessarily true. This allows for the possibility of finding informative descriptions of the essence of an individual, in the sense that such a description would be true of the individual no matter how the universe might have been. Moreover, the correct semantic theory allows us to speak meaningfully of counterfactual situations in which inessential descriptions are negated; this allows us to specify which actually-true descriptions pick out inessential features of individuals.

With the significance of these distinctions established, Kripke goes on to articulate a number of metaphysical and epistemological conjectures. One of them is the

necessity of identity. Kripke's example is "Hesperus is Phosphorous". Given rigid designation, the name "Hesperus" refers to the same object—Venus—in all possible worlds, as does the name "Phosphorus". Therefore, there is no possible world in which "Hesperus is Phosphorus" is false, hence that sentence expresses a necessary truth. Yet we can grasp the semantic content of these names without knowing this necessary truth; to learn that Hesperus is Phosphorus requires astronomical observation. Hence it is a necessary but a posteriori truth.

The companion case is one of contingent a priori truth. When we reflect on the sentence "The meter bar is one meter long" we can know that it is true just by knowing the semantic properties of the expressions "The meter bar" and "is one meter long". It takes no observations to know the truth of this sentence; hence it is a priori knowable. But because we know that "is one meter long" is not necessarily coextensive with "The meter bar", we also know that the sentence is contingent.

Another conjecture is the essentiality of origins. Kripke considers the Queen of England and asks whether she might have had different parents. He suggests that there might have been a person born of different parents who had all of the Queen's describable properties—even her genetic code, if we allow for some very fortuitous genetic mutation—but nevertheless Kripke maintains that we would not wish to identify the offspring of different parents as the very same being as the Queen. The argument consists in judiciously deployed italics: "How could a person originating from different parents, from a totally different sperm and egg, be *this very woman*?" Kripke follows up this argument with the following additional considerations: "it's possible that even

though she were born of [her own] parents...she never became Queen...[having been switched with another child at birth.] But what is harder to imagine is her being born of different parents. It seems to me that anything coming from a different origin would not be this object."

Along these same lines, Kripke suggests that substances that can be identified by name, such as gold or water, can also be identified by descriptions that are necessarily coextensive across possible worlds with the names that rigidly designate them. This allows sentence to express necessary truths about essence, such as "Gold is composed of atoms composed of 79 neutrons" and "Water is composed of H_2O molecules". These essence statements can be established by empirical means, but the fact that they *are* statements of essence—and therefore hold of necessity—is established by reflecting on semantics.

What is striking about each of these examples, from the perspective of Neurathian physicalism, is that they are established on the grounds of rational insight into the nature of semantic properties. From a syntactic point of view, reflection on the use of language cannot establish any conclusions about what is necessary or what is essential to an individual or a substance. The program of unified science might predict that the terms "Hesperus" and "Phosphorus" will be used interchangeably for the foreseeable future, but no conclusion could be drawn about the necessity of identity because syntacticism admits of no notion of semantic value across possible worlds. In the case of "The meter bar is one meter long", if we disavow all knowledge of semantics then we cannot claim to know

that “The meter bar” refers to the meter bar in the possible world where the meter bar is less than a meter long; hence, we cannot claim to know that sentence is contingent.

The question of the essentiality of the Queen’s origin turns on whether the description “born from parent A and parent B” is necessarily coextensive with the Queen’s name. The question of the identity of water with H_2O similarly depends on the modal semantics for the word “water” and the description “H_2O”. These questions cannot be answered from the point of view of Neurath’s physicalism; there are no answers to semantic questions.

Naming and Necessity is the most cited work of analytic philosophy, and its conclusions (only a handful of which have been touched on here) are based entirely on speculation and rational insight concerning semantics. It is perhaps the most influential piece of 20th century philosophy, and its influence was enabled by the semantic turn. If Neurath had succeeded in his sociological program, Kripke’s metaphysical insights might not have had so much traction and a great deal of metaphysics perhaps would not have occurred.

It should be noted that the program of unified science is not opposed to incorporating modal logic as formalism. Even models of possible worlds may be incorporated, so long as these models are treated as syntactic entities—that is, as physical structures that exist within our space-time. Allowing modal logic does not commit the Neurathian physicalist to necessary and contingent truth as long as it is treated syntactically.

6.3 Semantic Externalism and the Metaphysics of Correspondence

Putnam (1975) initiates another major metaphysical program of the 20th century through semantic argumentation. In his much discussed Twin Earth thought experiment, readers are asked to imagine that there is a planet that is identical to the Earth in every way except that the chemical composition of the substance comprising the twin planet's oceans, rivers, bathtubs, etc. is not water but another chemical substance abbreviated "XYZ". Putnam then asks us to consider the semantic content of the expression "water" as used on Twin Earth. The proper semantic insight to be drawn according to Putnam is: on Twin Earth the word "water" designates XYZ, not water. Yet the people on Twin Earth are physically and functionally the same as we are in all relevant respects. The conclusion is that the semantic content of linguistic expressions does not supervene on the internal physical nature or structure of an organism, but rather on relations between the organism and its environment. Concordantly, the truth of sentences is determined in part by such relations. The sentence "There is water on Earth" is true but the sentence "There is water on Twin Earth" is false, in part because there is no XYZ on Earth for our term "water" to refer to.

The metaphysical fallout of this semantic argument is immense. Our understanding of scientific truth must take into account not just the facts that we investigate, but also the relationships between the organism and its environment that constitute the semantic-content side of the correspondence between our theories and reality. Moreover, these representation relationships can hold without our knowledge. Thus our 'operational definitions' by which we control the use of our terms—our best

descriptions of the nature of physical substances and structures around us—can turn out to be false given the real essences of the substances that our words actually refer to in virtue of this external reference-fixing relation. This provides for the possibility that even our best possible scientific theories may be false; it also provides the possibility that some description available in our language or some more sophisticated language might be absolutely true, in the sense that the descriptions made available by the ideal language would capture reality exactly as it corresponds to the language we now use.

This type of metaphysicalized philosophy of science surely has Neurath spinning in his grave. The notion that facts about reality that are beyond the grasp of even our best theories might determine the absolute truth or falsity of any given encyclopedia is perhaps the farthest antithesis of Neurathian physicalism.

The syntacticist would reject Putnam's earliest premise. There is no sense to be made, from a syntactic perspective, of the claim that "water" refers to XYZ as it is used on Twin Earth; nor would the syntacticist concede that "water" refers to water as it is used on Earth. The program of unified science would be capable of describing the patterns with which the term is used on Twin Earth and contrasting it with the way it is used on Earth, but these differences would be superficial. Among other pedestrian predictions, the physicalist would expect that Twin Earthlings would use their word "water" in a similar way to the way Earthlings use the word "water", but nothing of metaphysical significance would be concluded from this prediction. Certainly nothing about the nature of reality or the nature of scientific inquiry could be concluded from

similarities or differences in the deployment of language construed as mere physical structure.

6.4 Mental Content Externalism

Invoking semantic intuitions similar to those used by Putnam, Burge (1979) presents a semantic argument for the claim that mental states supervene on social linguistic practices.

Burge gives the following thought experiment. First, imagine that there is an individual who sincerely tells his doctor "I have arthritis in my thigh." We are to imagine that the patient has a number of additional beliefs he would express with sentences using the word "arthritis" such as "I have arthritis in my wrists", "My father had arthritis", and "Arthritis is not worse than cancer," all of which his doctor would assent to. But of course his doctor says, as we would, that the belief he expresses with the sentence "I have arthritis in my thigh" is false. Next we are to imagine a counterfactual scenario in which the patient is exactly physically the same, with the same ailments, but in which his community uses the word "arthritis" to refer to the ailment the patient in fact has in his thigh. In this scenario, the patient's doctor says that the patient is correct.

Burge argues as follows. First, we are invited to agree that the sentence "I have arthritis in my thigh" is true in the counterfactual scenario, where it was false in the imagined actual scenario. Second, we are invited to agree that the sentence uttered expresses the patient's honest belief in both the actual and counterfactual scenarios. From these two premises, it follows that the patient expressed a different belief in the

counterfactual scenario from the one he expressed in the imagined actual scenario. The conclusion Burge draws is that belief supervenes on socio-linguistic practice and not on any intrinsic physical properties of individuals.

Mental externalism of this kind (and it can also be derived from a modification of Putnam's externalist argument) has generated a large literature on the metaphysics of mental states. All of the positions adduced in that literature depend for their support on rational insight concerning the semantic properties of mental states and the linguistic constructions that express and attribute them. From the perspective of syntacticism, these metaphysical debates and speculations cannot get off the ground. The intuitive judgments about sameness and difference in semantic content are properly regarded as pseudo-statements from the perspective of Neurathian physicalism. Again, unified science can account for sociological differences in patterns of linguistic usage construed as physical phenomena, including linguistic constructions used to classify mental states, but no substantive metaphysical conclusions can be reached from these patterns of linguistic usage.

6.5 States of Affairs and Truthmakers

According to Armstrong (1993, 1997), a state of affairs exists if and only if a particular has a property, or a relation holds between two or more particulars. Moreover, states of affairs are constituted by the particulars, properties, and relations that are sufficient for their existence. Armstrong takes properties and relations to be universals.

The argument for states of affairs is semantic. Armstrong writes, "Why is anything more needed than particulars and universals, monadic and polyadic? The answer to this comes from one of the fundamental assumptions that drive this ontology. It is the need for truths to have a truthmaker...or an ontological ground..." The idea is that the mere existence of e.g. a property F and a particular *a* is not sufficient to guarantee the truth of "*a* is F". If the sentence is true, then (according to Armstrong) some truthmaker is required. "The state of affairs of *a's being F* is suggested as that truthmaker, as the ontological ground." Hence the acceptance of a semantic notion of truth engenders discussion of states of affairs, a type of entity posited as the ground of truth, as well as speculation concerning the reality of properties, relations, and particulars, as the things that constitute states of affairs.

The notion of a state of affairs brings with it a host of metaphysical puzzles. One example is the question of negation. What is the truthmaker for the sentence "a is not F"? Does the negation sign designate a property of negation? Are there negative states of affairs? Are properties also particulars? If they are not, then how can we name them? Armstrong's picture also gives rise to the challenge of Bradley's regress. Armstrong raises the question this way: "Is not bringing the constituents of a state of affairs, the particulars, the properties and the relations, together into states of affairs, a further relation in which all the constituents stand? But then the new relation is just a further element which requires to be integrated along with the other constituents." If being part of a state of affairs is itself a state of affairs, then each state of affairs requires infinitely many more states of affairs to hold it together.

Neurath's syntacticism undermines Armstrong's discussion of states of affairs in a most straightforward way. Since sentences do not have properties such as truth and falsehood, there is no need to explain the truth of "a is F" by positing some ontological structure in which the truth of the sentence is grounded. More generally, the metaphysics of truthmakers and truthmaking is rendered completely idle by the syntactic perspective.

6.6 Concluding remarks

I have presented only a very small sample of all the possible examples that attest to the many ways in which the semantic turn has given rise to metaphysics. The list of such case studies could be expanded indefinitely as long as the current philosophical era persists. I take the point to be proven that Neurath was right about the danger that semantics poses to logical empiricism. It was an empirical question whether Carnap was right to think that a semantic theory of truth could be added harmlessly to the scientific worldview. The evidence strongly suggests that he was wrong. A physical syntacticism of the kind advocated by Neurath would have provided a strong counter-position to the metaphysical discourse of the twentieth century if it could have been maintained. Perhaps that is why Neurath's opposition to a semantic theory of truth was so vehemently rejected.

References

Armstrong, D.M. (1993) A World of States of Affairs. Philosophical Perspectives, Vol. 7, Issue Language and Logic, 429-440

Beth, E. W. (1963) "Carnap's Views on the Advantages of Constructed Systems Over Natural Languages in the Philosophy of Science." In Schilpp (1963) The Philosophy of Rudolf Carnap.

Burge, Tyler (1979). Individualism and the mental. Midwest Studies in Philosophy 4 (1):73-122.

Carnap, Rudolf (1922). Der Raum: Ein Beitrag zur Wissenschaftslehre. Kant-Studien 56.

---- (1934/1937). The Logical Syntax of Language. London, K. Paul, Trench, Trubner & Co., Ltd.

---- (1950). Empiricism, semantics, and ontology. Revue Internationale de Philosophie 4 (2):20--40.

---- (1974). An Introduction to the Philosophy of Science. Dover.

Cartwright, Nancy ; Cat, Jordi ; Fleck, Lola & Uebel, Thomas E. (1996). Otto Neurath: Philosophy Between Science and Politics. Cambridge University Press.

Creath, Richard (1990). The Unimportance of Semantics. PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association 1990:405 - 416.

Coffa, Alberto (1987). Carnap, Tarski and the search for truth. Noûs 21 (4):547-572.

Field, Hartry (1994). Deflationist views of meaning and content. Mind 103 (411):249-285.

Friedman, Michael (1999). _Reconsidering Logical Positivism_. Cambridge University Press.

Hempel, Carl G. (1935). On the Logical Positivists' Theory of Truth. _Analysis_ 2 (4):49 - 59.

Hill, Christopher S. (2002). Thought and World: An Austere Portrayal of Truth, Reference, and Semantic Correspondence. Cambridge University Press.

Horwich, Paul (1998). Truth. Clarendon Press.

Joachim, Harold H. (1906). The Nature of Truth. New York, Greenwood Press.

Mancosu, Paolo (2008). Tarski, neurath, and kokoszynska on the semantic conception of truth. In Douglas Patterson (ed.), New Essays on Tarski and Philosophy. Oxford University Press 192.

David, Marian (1994). Correspondence and Disquotation: An Essay on the Nature of Truth. Oxford University Press.

Mormann, Thomas (1999). Neurath's Opposition To Tarskian Semantics. Vienna Circle Institute Yearbook (6):165-178.

Nagel, Ernest (1939). The Formation of Modern Conceptions of Formal Logic in the Development of Geometry. Osiris, vol. 7 pp.142-223

Neurath, O. (1913) The Lost Wanderers of Descartes and the Auxiliary Motive. In Neurath 1983, p. 1-13.

---- (1915) On the Foundations of the Histoy of Optics. *Archiv fur die Geschichte der Naturwissenschaften und der Technik*, Bd. 5, 371-389. Leipzig, F. C. W. Vogel. In Neurath 1973, p. 101-112.

---- (1916) On the Classification of Systems of Hypotheses (With Special Reference to Optics). *Jahrb. D. phil. Gessell. a. d. Univ. Wien.* In Neurath 1983, pp.13-32.

---- (1921) Anti-Spengler. In Neurath, O.; Neurath, Marie & Cohen, Robert S. *Empiricism and Sociology*, p. 158-213.

---- (1930) Ways of the Scientific World Conception. Erkenntnis 1, 106-125. In Neurath 1983, p. 32-48.

---- (1931a) Physicalism: The Philosophy of the Viennese Circle. The Monist 41 (4), 618-623. In Neurath 1983, p. 48-52

---- (1931b) Physicalism. Scientia (50): 297-303.

---- (1931c) Empiricism and Sociology. Vienna: Springer. In Neurath 1973, p. 319-420.

---- (1931d) Sociology in the Framework of Physicalism. Erkenntnis 2 (1):393-431. In Neurath 1983 p. 58-91

---- (1934). Radical Physicalism and the 'Real World'. Erkenntnis 4 (1):346-362. In Neurath 1983 p. 58-91.

---- (1936d) Encyclopedia as 'Model'. Revue de Synthese 12, 187-201. In Neurath 1983, p. 145-159.

---- (1936e) Physicalism and the Investigation of Knowledge. In Neurath 1983, p. 159-172.

---- (1983) Philosophical Papers 1913-1946. Dordrecht, Holland. D. Reidel Publoshing Company.

Neurath, O.; Neurath, Marie & Cohen, Robert S. (1974). Empiricism and Sociology. Dordrecht, Holland. D. Reidel Publishing Company.

Noonan, J. (1990). Carnap's *Logical Syntax of Language*. Unpublished senior thesis, Department of Philosophy, Boston University.

Oberdan, Thomas (1992). The concept of truth in carnap'slogical syntax of language. Synthese 93 (1-2):239 - 260.

Putnam, Hilary (1975). The meaning of 'meaning'. Minnesota Studies in the Philosophy of Science 7:131-193.

Sarkar, Sahotra (1992). "The boundless ocean of unlimited possibilities": Logic in Carnap's logical syntax of language. Synthese 93 (1-2):191 - 237.

---- (2013). Carnap and the compulsions of interpretation: Reining in the liberalization of empiricism. European Journal for Philosophy of Science 3 (3):353-372.

Schilpp, Paul Arthur (ed.) (1963). The Philosophy of Rudolf Carnap. La Salle, Ill.,Open Court.

Schlick, Moritz (1930). Die wende der philosophie. Erkenntnis 1 (1):4-11.

---- (1934). Über das Fundament der Erkenntnis. Erkenntnis 4 (1):79-99.

Gödel, K. (1995). Is Mathematics the Syntax of Language? In K. Gödel (Ed.), Kurt Gödel: Collected Works. Vol. 3 (pp. 334–362). Oxford: Oxford University Press.

Uebel, T (1996). On Neurath's Boat. In Otto Neurath: Philosophy Between Science and Politics. Cartwright (1996) Cambridge University Press.

---- (2001). Carnap and Neurath in exile: Can their disputes be resolved? International Studies in the Philosophy of Science 15 (2):211 – 220.

---- (2007) *Empiricism at the Crossroads: The Vienna Circle's Protocal-Sentence Debate Revisited*. Open Court.

---- (2009). Neurath's protocol statements revisited: sketch of a theory of scientific testimony. Studies in History and Philosophy of Science Part A 40 (1):4-13.