

Frege on the Indefinability of Truth

Hans Sluga

The content of the word ‘true’ is altogether unique and indefinable”—so Frege in his 1918 essay “The Thought.”¹ The provocative remark has caused his interpreters some embarrassment, for shortly afterward Alfred Tarski was to show how truth is definable in certain formalized languages, and this result is generally now considered one of the major achievements of modern logic. As a result, contemporary critics consider Frege’s thesis commonly as out of date or as ill conceived from the start. They also note that he proposes it only once in his published work, at the beginning of “The Thought,” and that its bearing on the rest of this essay, not to speak of the rest of his work, is far from evident. They express bewilderment, moreover, at the sketchy argument “The Thought” advances in its support.

The easy way out would therefore be to ignore Frege’s claim as an unfortunate blunder. But further scrutiny reveals that he himself considered it crucially important and that it lies behind much else he wrote about truth. What is more, Frege’s thesis contains *philosophical* insights into the notion of truth that are not captured in Tarski’s work and that have generally been ignored by those intrigued with Tarski’s technical brilliance. A popular misreading of Tarski has it that his semantics provides the key to an understanding of the ontological structure of the world. The unfortunate result is a revival of an a priori form of metaphysics that we had thought ourselves well rid off. Proper reflection on Frege’s thesis can serve as a timely antidote to such misconceptions and is, therefore, of more than historical interest.

I. The Place of “The Thought” in Frege’s Work

In order to see what significance attaches to the indefinability thesis, we must first be clear about the special place it occupies in “The Thought” and the equally

special place of the essay itself in Frege's oeuvre. For "The Thought" was by no means an incidental piece of writing. Frege published it in 1918, at the age of seventy, as the first in a series of articles under the general title called "Logical Investigations." It was followed in the same year by a second piece called "Negation" and in 1923, two years before his death, by a third entitled "Thought Connections." His posthumous writings contain, in addition, the beginning pages of an article on "Logical Generality" that was clearly meant to form a fourth installment of the "Logical Investigations." And there are indications that he intended to continue the series even further. Notes written in 1919 for the historian of science Ludwig Darmstaedter outline a course of thought that leads from the content of the four pieces just mentioned to an account of the sense-reference distinction and on to reflections about the nature of arithmetic and the numbers. It seems plausible to conclude that Frege planned to extend his investigations in this direction. This would also make sense of a number of drafts from the last two years of his life in which he sketches an altogether new approach to the analysis of the numbers.

In composing "The Thought" and the essays that followed, Frege drew on material of an earlier date. Since the 1880s he had been at work on an informal account of his logical and philosophical doctrines. In a series of drafts, with such titles as "Logic," "Introduction to Logic," "What Can I Regard as the Result of My Work?," "Short Survey of My Logical Doctrines," and "My Basic Logical Insights," he had tried to expound his ideas in a number of ways without, however, managing to bring them into a publishable form. The 1918 essay and its sequels were, in fact, based on the most extensive of these pieces, a text called "Logic" from 1897. Frege was, therefore, right when he suggested to Wittgenstein in 1918 that there was perhaps "little that is new" in "The Thought." But he was also right when he added that "perhaps it is still said in a new way and thus more comprehensible for some."² For while "The Thought" follows the earlier draft in both conception and language, there remain significant differences between the two pieces, and these are often as interesting as their agreements.

There is no doubt, in any case, that the composition and publication of "The Thought" constituted a surprising new departure for Frege. Disheartened by the inconsistency in his formal system, he had published nothing in the preceding decade. He was retiring from his university that year after forty-four years of teaching and was leaving the city of Jena, his home for half a century, to move to the tiny and isolated town of Bad Kleinen. He had not been in good health for some while, and Germany's military, political, and economic situation clearly depressed him. Given that the moment was in every respect inauspicious, one is left to wonder what may have helped to renew Frege's intellectual energies at this time.

The explanation that suggests itself most strongly is his relation with Wittgenstein. In 1911 he had come into contact with Wittgenstein, who had treated him and his work with unusual consideration. Wittgenstein visited him in 1911, 1912, and 1913 to talk extensively about logic and philosophy, with each of these visits lasting several days. During the war the two maintained contact, though little of their correspondence appears to have been devoted to philosophical issues. Still, Frege knew at least that Wittgenstein was continuing his work even under the most daunting circumstances, and his letters intimate that this reinvigorated his own intellectual energies. Hearing in May 1918 that Wittgenstein was finally putting

his thought into writing, he responded: "Perhaps, I too will be helped by this in the difficult territory in which I am laboring. I always expect some profit from getting to know the ways you have traveled, even when I should not be able to follow you in the essentials" (B, p. 17). One may plausibly conjecture, then, that it was the prospect of Wittgenstein's treatise that spurred Frege on to resume his labors. The supposition is further confirmed by his reaction to the news, four months later, that Wittgenstein had actually completed his work. His response combines congratulations with the announcement that he would soon be sending a copy of an article of his own, that is, "The Thought." His letter also predicts (correctly, as it turned out) that Wittgenstein would "probably not completely agree" with it but voices the hope that it would lead to lively exchanges between them.

Frege's new venture had been made possible by the philosopher Bruno Bauch, who at that time offered to publish something by Frege in his *Beiträge zur Philosophie des deutschen Idealismus*. This was a novel experience for Frege, who had always encountered difficulties in getting his articles printed. Bauch had founded his journal at the end of World War I as a nationalistic alternative to the liberal *Kantstudien*. To assure its success, he had sought to attract distinguished older academics sympathetic to its political line—among them the psychologist Wilhelm Wundt and his former teacher Heinrich Rickert, who had both veered to the right in World War I. As one of his colleagues at Jena, Bauch was presumably also familiar with Frege's increasingly reactionary views and was evidently keen to draw him as well to the journal and its associated organization, the Deutsche Philosophische Gesellschaft. He had, moreover, a genuine interest in Frege's work and had endeavored to use it in his own philosophy—a combination of Kantian idealism with a theory of objective value due to Lotze. Frege had his own obligations to Lotze and had borrowed from him, in particular, the distinction between subjective mental representations and objective thought. His essay "The Thought" concerned, among other things, precisely this Lotzean theme. When it appeared in the first volume of the *Beiträge*, Bauch preceded it appropriately with a piece of his own on Lotze's logic—clearly intended to legitimate the connection between Frege, Lotze, and himself. But Frege's intentions in publishing that essay went in a different direction. For Bauch's generous offer provided him with the unexpected chance to give his views a final and definitive airing before Wittgenstein could lay out his related though distinct ideas.³

Only if we are clear about this prehistory are we likely to appreciate the weight carried by the initial pages of "The Thought." Frege begins that essay with the declaration that logic is fundamentally concerned with truth: "As the word 'beautiful' points the direction for aesthetics and 'good' that for ethics, so 'true' points the direction for logic. . . . To discover truths is the task of all sciences; logic is concerned with recognizing the laws of truth" (T, p. 17). He goes on to denounce the idea that logic might be speaking of the mental process of thinking; the domain of logic is rather that of objective thought and its attendant truth or falsity. He insists on a boundary between psychology and logic that must not be breached, and he proceeds from this to drawing "the rough outlines of what I want to call true in this context."

Truth, he points out, is commonly asserted of "pictures, representations, propositions, and thoughts," and he considers the possibility that in proposi-

tions as in pictures we speak of truth “insofar as there exists a correspondence [*Übereinstimmung*] between the picture and what it depicts.” But a correspondence, he insists, “can only be perfect if the corresponding things coincide and are, therefore, not distinct things at all.” But this is not what we expect in the case of propositions. Here we assume, at most, that idea and reality correspond in certain respects. However, in this case, “we should have to inquire whether it were true that an idea and reality, perhaps, corresponded in the laid-down respect. And then we should be confronted by a question of the same kind and the game would begin again.” He proceeds from this to a generalized attack on all attempts to define the notion of truth, arguing that for the same reasons as the correspondence theory of truth “every other attempt to define truth collapses too.” Any definition whatever of the concept of truth would have to specify that a proposition is true, if and only if it has certain characteristics. But in this case “the question would always arise whether it were true that the characteristics were present. So one goes round in a circle.” The ultimate conclusion is, then, that “the content of the word ‘true’ is unique and undefinable” (T, pp. 18–19).

II. Six Phases in Frege’s Thinking about Truth

This statement of the undefinability thesis marks the terminal point in the evolution of Frege’s thinking about the notion of truth. The process leading to it had extended over forty years and had gone through at least seven distinctively different phases. Mapping these phases is helpful for understanding the important function that Frege assigned to his thesis. It will also help to destroy the common belief that Frege was an essentially static thinker whose system of ideas had come to him in one piece. Nothing reveals the dynamics of Frege’s thought more clearly than following the track of his thinking about truth from 1878 to 1918.

Phase I: A Logic of Judgment Rather Than Truth (1878)

The first phase of Frege’s thinking about truth is marked by the *Begriffsschrift*, which he finished in late 1878. In this phase the notion of truth plays a peculiarly subsidiary role. Frege’s explicit references to it are almost all confined to the first paragraph of the preface of his monograph. After that point the notion is set aside in favor of other concepts. What is distinctive in this manifests itself already in the very first words of the *Begriffsschrift* when he writes: “In apprehending a scientific truth we pass, as a rule, through various degrees of certainty.”⁴ Though there is talk here of truth, one is struck by the fact that Frege’s attention is directed toward its apprehension rather than truth itself. And this focus is maintained in the next few sentences, where he speaks of propositions coming to be more securely established “by being connected with other truths through chains of inferences” and of the division of “all truths that require justification into two kinds,” those that can be proved “purely by means of logic” and those whose proof requires appeal to “facts of experience.” In each of these claims Frege speaks *cognitively* of “recognition,” “certitude,” “justification,” and “secure foundations,” of propositions more or less “securely established,” “derived,” or “confirmed,” “in need of justification” or not, “supported by facts of experience” or depending “solely

on those laws on which all knowledge rests." He is evidently preoccupied throughout this passage with questions of knowledge rather than with truth per se (BS, pp. 5–6).

This is immediately reflected in the logic he proceeds to construct. For it is not a logic of truth and falsity but one of judgment. What he will later call a "thought," that is, the sense of a declarative sentence, he calls here a "judgeable content." And the primary symbol of the new logic is the "judgment sign," \vdash , which has the double function of combining the signs that follow it into a whole and of "affirming" this whole (BS, p. 12). This sign is so fundamental that Frege allows that one might consider it the single and common predicate of all judgments expressible in his symbolism: "We can imagine a language in which the proposition 'Archimedes perished at the capture of Syracuse' would be expressed thus: 'The violent death of Archimedes at the capture of Syracuse is a fact.' . . . Such a language would have only a single predicate for all judgments, namely, 'is a fact.' . . . Our *Begriffsschrift* is a language of this sort, and in it the sign \vdash is the common predicate for all judgments" (BS, p. 12f.).

The second logical symbol of the *Begriffsschrift*, that for the conditional, is also introduced without reference to the notion of truth. In characterizing the conditional judgment, Frege makes use instead of the notions of affirmation and denial. He writes:

If A and B stand for contents that can be judged, there are the following four possibilities:

- (1) A is affirmed and B is affirmed;
- (2) A is affirmed and B is denied;
- (3) A is denied and B is affirmed;
- (4) A is denied and B is denied.

Now $\begin{array}{|c} \vdash & A \\ & | \\ & \vdash & B \end{array}$

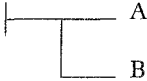
stands for the judgment that the third of these possibilities does not take place, but one of the others does. (BS, p. 13f.)

The first point to note is that he does not here explain the meaning of a connective but that of a conditional judgment. Equally important is that he characterizes the conditional by the analogue of a truth-table but that his operative terms are those of affirmation and denial, not those of truth and falsity. Instead of saying that the conditional sentence is true or false depending on the truth-values of its constituents, Frege says that the conditional judgment stands for some possibilities of affirmation and denial taking place and others not.

This explains also the peculiarities of the notation he adopts. We tend to think of a sentence of the form "If B, then A" as expressing a relation between the "antecedent" B and the "consequent" A. Frege conceives of it, instead, as an affirmation of A that differs from the straightforward, unconditioned affirmation

$\vdash A$

only in that it is made on the basis of another affirmation. That the affirmation of A rests conditionally on B, Frege's notation indicates by letting the symbolic expression of the affirmation of A rest literally on B. Hence,



Though Frege speaks of logic throughout the *Begriffsschrift* in terms of affirmation and denial, this should not be taken to mean that truth plays no role whatsoever in his account. His intentions can be made clearer if we distinguish between implicit and explicit notions within his logic. The introductory sentences of the *Begriffsschrift* show him to be aware of the notion of truth as essential to logic, but he reserves for it an implicit role. The basic *explicit* term is that of judgment. But to judge that A is always to judge that A is true and there is, hence, for Frege an internal link between the notions of judgment and truth.

Phase II: The Logic of Right and Wrong (1881)

In three closely related papers written in 1881–82 Frege sets out to reply to criticisms Ernst Schroeder had leveled against the *Begriffsschrift*.⁵ In defending himself in these papers, entitled respectively “Boole’s Calculating Logic and the *Begriffsschrift*,” “Boole’s Logical Formula Language and My *Begriffsschrift*,” and “On the Aim of a *Begriffsschrift*,” he takes the opportunity to restate the fundamental assumptions and concepts of his logic. This he does, in part, by using his old language of affirmation and denial. Thus, he explains a formula we would render as

$$(P \ \& \ \neg \ Q) \rightarrow R$$

as saying that “if P is affirmed and Q is denied, then R is affirmed.”⁶ But in the general characterization of the conditional itself he takes an altogether new route. The first indication of this is that he now defines the conditional *connective* and not, as he had previously done, the conditional judgment. And in introducing that connective he no longer employs the language of affirmation and denial but says, instead, that the conditional sign expresses “the negation of the case that the upper content is wrong [*falsch*] and the lower is right [*richtig*]” (NS, p. 12). From this follows, even more important, a wholly new characterization of logic itself. While the *Begriffsschrift* had maintained the traditional characterization of logic as a concern with the laws of thought, Frege now has it that logic deals with the laws of “right inferring” [*richtiges Schliessen*] (NS, p. 13).

The move is certainly a technical improvement over the *Begriffsschrift* account, which had suffered from at least two flaws. The first was that Frege had introduced the conditional there by speaking of “possibilities” of affirmation and denial when he had at the same time dismissed modality as being “without meaning” for his logic (BS, p. 13). The second flaw lay in the *Begriffsschrift*’s confusion between whole judgments and their judgeable parts. Thus, Frege had said of the conditional judgment as a whole that it *affirms* the conditional “If B, then A,” but he had also said that both its constituents “A” and “B” are affirmed or denied. He had, thus, used the term “affirmation” to characterize both the distinctive function of the judgment as such *and* the contribution that its components make to the judgment. The new terminology of 1881–82 avoided such pitfalls.

The new terminology signals thereby a turn to a decidedly more “objective” specification of logic. For right and wrong are, presumably, judgment-independent characteristics of propositions. But the question remains why Frege introduced a wholly new pair of terms to make that point. Why did he not fall back on the notions of truth and falsity, which had been available to him since the *Begriffsschrift*? The answer may be this. Both the German terms *richtig* and *falsch* and their English equivalents “right” and “wrong” contain a reference to an outside ruler or standard. *Richtig* and “right” mean originally straight and aligned with a ruler and then, more generally, in accordance with some standard or measure. *Falsch* and “wrong” mean not in accordance with such a standard. The new terms have, thus, an inherently relational meaning. Right inferring is, on this understanding, an inferring according to logical laws, and in this lies for Frege the kernel of an important thought he will pursue later on. What turns out to be more problematic for him is the question of how we are to understand talk of right and wrong propositions. We know that Frege later rejected the idea that truth is relational in character. But the terms “right” and “wrong” have an inherently relational meaning. He does not actually tell us in these writings what standard or measure a right proposition is meant to accord to. But it is not absurd to conjecture that at this point a proposition is right for him when it corresponds to reality, that reality is the standard to which propositions need to be aligned.

If so, Frege never gets to spell this out. The second phase in his thinking about truth proves, instead, merely transitional. He soon abandons the terminology of right and wrong and replaces it by talk about true and false. But if our conjecture is correct, this apparently innocuous shift hides an important move in Frege’s thinking—a turning away from the idea that truth is to be understood as a relational notion and with it, a fortiori, a turning away from the idea that truth can be understood as a correspondence to facts. The seeds of Frege’s later explicit rejection of the correspondence of truth may, thus, be found in his abandonment of the terminology of right and wrong that he had entertained in 1881–82.

Phase III: The Objectivity of Truth (1884)

Truth turns into a central notion finally in the *Foundations of Arithmetic* of 1884. Frege now characterizes logical proof directly in terms of “truth.” A proof, he writes, seeks “to place the truth of a proposition beyond all doubt,” and it also “affords us insight into the dependence of truths upon one another.” He worries now over the question of whether “arithmetical truths” are synthetic or analytic, and he speaks broadly of “truths of a general logical nature,” of truths that “belong to the sphere of some special science,” of “truths which cannot be proved,” and of truths a priori and truths a posteriori. Most important, there are for him now “primitive truths” (*Urwahrheiten*) to which our inquiries will eventually reduce everything.⁷

It is legitimate to ask what motivated Frege to adopt that new language at this point. The notion of truth had remained implicit and hidden in the *Begriffsschrift* period and had been replaced by that of rightness in the next phase of his development. But now, in this third period, it seems to proliferate. What, we may ask, brought about this change of mind? We have noted already that the shift to the notions of right and wrong marks an increasing preoccupation in Frege’s mind

with the objectivity of logic. That concern becomes even stronger in the third phase and coupled with the rejection of any relational account of the rightness of propositions leads now to a new emphasis on the notion of truth.

Frege's increasing preoccupation with the objectivity of logic is made evident, in the methodological principles he formulates for himself at the beginning of *The Foundations of Arithmetic*. The first of these is "always to separate sharply the psychological from the logical, the subjective from the objective" (F, p. x). This had not been his main concern in the period of the *Begriffsschrift*, where he had spoken in a broadly psychological language of ideas and their combination. In his essay "On the Scientific Justification of a *Begriffsschrift*" from the same period, he had talked of sense impressions and memory images as fundamental to human thought and had explained the advantage of linguistic symbols by arguing that they create a new focus around which a group of memory images can gather and, thereby, give rise to thinking in concepts. "Thus, in applying the same symbol to different, but similar things, we actually no longer symbolize the individual thing, but rather what they have in common: the concept."⁸ By 1884, however, Frege has turned staunchly antipsychological in his thinking. John Stuart Mill's *Logic* is now at the center of his attention, and in opposition to Mill he insists now on an absolute separation of the logical from the psychological. In *The Foundations of Arithmetic* he protests accordingly against a "preponderance of the psychological point of view in philosophy, which has even made incursions into logic," and explains in this fashion the reluctance of mathematicians to address philosophical questions, and, to counteract these tendencies, he emphasizes the distinction between the being thought of a proposition and its truth. "One must, so it seems, call to mind once again the fact that a sentence ceases as little to be true when I am not thinking of it, as the sun is destroyed when I close my eyes." Frege's turn to the concept of truth proves thus to be a direct result of his increasingly vociferous antipsychologism (F, pp. 6, 7).

That it is antipsychologism which motivates the recasting of his logic in the language of truth and falsity is confirmed by another text from the same period, a fragmentary piece, called simply "Logic," that contains the earliest of the already mentioned informal accounts of Frege's logic. The editors of his *Posthumous Writings* date the piece somewhere between 1879 and 1891. But we can say with some assurance that it must have been composed immediately before *The Foundations of Arithmetic*. It was certainly written after 1882, since the claims that "the aim of scientific inquiry is truth," that the logical laws are "nothing but an explication of the content of the word 'true'," and that logic itself is concerned with the clarification of the notion of truth (NS, pp. 2-3), with their emphasis on the notion of truth, belong neither to the first nor to the second phase we have distinguished. But the text was just as certainly written shortly after the end of the second phase. For Frege still characterizes the laws of logic initially in terms of the language of the second phase as "laws of right inferring" but then corrects himself and calls them instead explications of the content of the word "true." This dating is confirmed by another piece of evidence, for the text is preceded by a table of its projected contents, which hints at the possibility of defining objects by means of "judgments of recognition." This idea is actually put forward in section 62 of *The Foundations of Arithmetic* but then rejected, and we are thus justified in concluding that the text must have been written shortly before 1884.

All this is of interest only because the fragment contains a number of ideas and formulations that will eventually recur in "The Thought." Both share with *The Foundations of Arithmetic* the belief that truth is to be distinguished from its recognition and that in being concerned with truth we are not dealing with something psychological. The business of the logician is, rather, "a continuous fight against the psychological." Frege also introduces here for the first time the idea that the notion of truth is to be associated with that of good in ethics. The logical laws are in some ways like moral principles, not like natural laws: they are laws "according to which right thinking proceeds," not laws of actual, psychological happenings. These formulations take up and expand on the idea voiced in the second phase that logic is concerned with the laws of right inferring. By this Frege had, presumably, meant that such laws are normative in character in that they adjudicate between right and wrong acts of inferring. After the second phase the idea that logic is concerned with inferring will drop away, but the belief in the normative character of the logical laws will remain. While Frege understands them no longer as (normative) laws of inferring, he now treats them as (normative) "laws of truth."

In accordance with this shift he writes now that "anyone who fails to recognize the unique meaning of the word [true] will be unable to get clear about the task of logic." While this comes close to his later views on truth, we must be careful not to read too much into these early formulations. For two important elements of Frege's fully matured view of truth are as yet absent. The first is the doctrine of truth-values, the second is an explicit statement of the thesis that truth is indefinable. For the first we will have to wait another six years, for the second another thirteen (NS, pp. 3–7).

Yet another thing is absent from both *The Foundations of Arithmetic* and the early "Logic" fragment. While both assign a central and explicit role to the notion of truth, neither of them makes the slightest attempt to define it. This is particularly striking in *The Foundations of Arithmetic*, where we are offered definitions of the notions of analytic and synthetic, a priori and a posteriori truth but not of truth itself. Neither text talks of truth in terms of the notions of correspondence, facts, or reality, and neither text entertains the possibility of any other kind of definition of truth. The reason for this may be found in the context principle announced in *The Foundations of Arithmetic* (F, p. x). For the claim that words have meaning only in the context of a sentence can be taken to imply that the meanings of sentence components must be explained by reference to the meanings of the sentences in which they occur. On a strong reading, this would seem to exclude any characterization of the truth of a sentence in terms of the meanings of its components, and with the rejection of any compositional account of truth, any definition of truth as correspondence would also be ruled out.

Phase IV: The Idea of Truth-Values

In the next phase Frege inaugurates an idea that is widely considered puzzling or even counterintuitive. It is the assumption of two truth-values, the True and the False—two objects denoted by declarative sentences. This unusual doctrine is clearly intended to serve two ends. It distances Frege from any account of truth as a property or relational characteristic of propositions. It also aims to account for

the normative character of the logical laws of which he had spoken previously. But his talk of sentences denoting objects has been almost universally dismissed as a conflation between the semantics of names and that of sentences and as explicable only by a false economizing with theoretical distinctions. Insofar as logicians still employ the language of truth-values, they do so today as a shorthand for what they take truth to be in reality, namely, a certain kind of correspondence and, hence, a relational property. It would be easy to ascribe to Frege such an attenuated doctrine, but his words point clearly in another direction.

Having come to think of the logical laws as specifically concerned with the notion of truth, the question becomes for him in this fourth phase one concerning the nature of their validity. While these laws are related to human thinking, they do not describe how thought actually proceeds but prescribe how it must proceed, if it is to attain truth. But why should truth and its pursuit matter to us? No account of truth as a natural quality of propositions can give us a satisfactory answer. For why should we be interested only in propositions that have that quality and not some other one? The objection holds obviously also for the correspondence conception of truth. Assume that some propositions correspond to reality and others not. Why should only the corresponding ones matter to us? It is as if propositions were divided into those written in black ink and those written in red, and we then expressed a preference for the former. This would seem to be at best an arbitrary choice. If human thought ought to aim at truth, then truth cannot be a natural property. There is hidden here in Frege's thought an argument equivalent to Moore's argumentation against naturalism in ethics.

Decisive for him is that truth must have a normative force and that its normativity must be objective. Frege finds himself here at a place which the Neo-Kantian philosopher Wilhelm Windelband had also reached at about the same time. Frege and Windelband had both studied with Kuno Fischer in Jena and with Hermann Lotze in Göttingen, and this may explain their convergent views concerning the nature of truth. It was Lotze who had introduced the language of value into philosophy and who had spoken of the Good and the Beautiful as objective values. In the early eighties Windelband had extended this doctrine to include the True as a value. He had argued that it was necessary to distinguish between the content of a judgment and its truth or falsity, and that the latter constituted "truth-values" (*Wahrheitswerte*) at which we aim in making a judgment.⁹ This still left the ontological status of such values uncertain. For Lotze they had been unreal (*unwirklich*) but objectively "valid" (*gültig*). Windelband and Heinrich Rickert, his student, eventually ascribed to them being and objecthood as a guarantee of their objectivity, and in this Frege seems to have followed them.

What separates him from the Neo-Kantians is, however, that he manages to integrate this speculative doctrine into the formal structure of his logic. The period in which he comes to speak of truth-values is also the one in which he works out his distinction between functions and objects and his classical theory of sense and reference. Ever since 1878 he had, in fact, operated with a distinction between functions and arguments meant to subsume the traditional philosophical divisions between subject and predicate, concepts and objects, but *generalizing* and re-describing them in mathematical terms. None of this had, however, been worked out in the early years. But by 1884 the matter was taking on a new urgency for him. He now thought of numbers as individuated objects rather than predicable

concepts and found it, thus, necessary to insist on the principle “never to lose sight of the distinction between concepts and objects” (F, p. x).

In further development of this distinction, he emphasizes now in the fourth phase that functions must be thought of as “unsaturated” or “incomplete,” and that they are as such *categorially distinct* from objects that are to be conceived as saturated or complete. Into this doctrine Frege inserts the idea of truth-values by identifying concepts with functions whose “value” for any argument is the True or the False. It follows from this that concepts are categorially distinct from objects and that a concept that can be predicated of an object cannot also be predicated meaningfully of a concept and vice versa. This creates a quandary for theoretical notions such as that of “being an object” or that of “being a concept.” We assume that they are legitimate and use them to classify and describe objects and concepts. But the doctrine of categorial distinctions proves this to be logically impossible. While we may have an intuitive grasp of the difference between objects and concepts, there are in Frege’s understanding no legitimate concepts at hand for a theoretical characterization of that difference. Anything that looks like such must be considered a practical device to be discarded once we have caught on to the distinction it seeks to describe.

From this follows that no theoretical account of the semantics of names and functional expressions is possible. For it would have to specify that names refer to objects but not to functions and that functional expressions refer to functions and not to objects. But this proves impossible if the predicates “being an object,” “being a concept,” “being a function” are categorially illicit. It follows, moreover, that even the notions of meaning and reference must be categorially illicit, if they are taken to apply to both names and functional expressions. And this leads almost immediately to the conclusion that truth must be indefinable—one which Frege will not actually state until later. But even in the fourth phase he is committed to the idea that the entities referred to by the components of a sentence are categorially distinct. In the simplest case, a sentence will have to contain a name standing for an object and a predicate standing for a concept. A truth definition would have to say in this case that the sentence is true if and only if the object named falls under the concept referred to by the predicate. It would require us to speak of “the meaning” of the sentence, the name, and the predicate, but on Frege’s account there is no single notion of meaning that could perform such a service. It follows that any definition of the concept of the truth of a sentence in terms of the meanings of its components is categorially defective.

Frege’s second important advance in the late 1880s is his doctrine of sense and reference. He comes to it through a rethinking of the account of identity he had given in the *Begriffsschrift*. In order to resolve the inadequacies of his earlier account, he insists now on a distinction between the sense and the reference of an expression. When he applies this distinction to a whole declarative sentence, he is moved to divide what he had previously called the judgeable content into that which the sentence expresses and the object to which it refers. Following Lotze he calls the former the “thought,” and following Windelband he calls the latter a truth-value. That this is formally satisfactory depends, of course, on the particular way in which he draws the distinction between sense and reference, and this is, in turn, interwoven with his account of concepts as truth-functions. It cannot be our goal here to describe the intricacies of this structure. It is enough to have

indicated that Frege's doctrine of the True and the False as values and objects referred to by declarative sentences is worked thoroughly into the formal machinery of his logic. Together, they lead him inevitably in the next phase to the explicit formulation of the thesis that truth is simple, unique, and indefinable.

Phase V: The Indefinability Thesis (1897)

The claim of the indefinability of truth is made explicitly for the first time in a reworking of the "Logic" fragment of 1884, which we can confidently date as having been written in 1897. Frege begins this text by drawing once more a comparison between good, beautiful, and true. While the first characterizes the "goal" of ethics and the second that of aesthetics, "true" characterizes the goal of science. Even though all science is concerned with truth, logic is concerned with it in a very specific manner. Like ethics, it can be called a normative science. Logic is "the science of the most general laws of being true" (NS, p. 139). Frege continues:

It would now be in vain to try to make clearer by a definition what is to be understood by "true." If one were to say: "An idea is true if it corresponds to reality," nothing would be gained, for in order to apply it we would have to ask in each case whether an idea corresponds to reality, in other words, whether it is true that the idea corresponds to reality. One would have to presuppose what is to be defined. The same would hold of any explanation of this form: "A is true if it has such and such a property or if it stands in this or that relation to this or that." . . . Truth is obviously something so primordial and simple that a reduction to something even simpler is impossible. We are therefore forced to illuminate what is unique in our predicate through a comparison with others. (NS, p. 140)

The first pages of "The Thought" will repeat this kind of reasoning, though in a somewhat more condensed fashion.¹⁰

Both pieces were, of course, written after Frege had introduced the sense-reference distinction, but they seem to pass over it with hardly a nod to the famous conception. There is, in particular, no mention here of the doctrine of truth-values. Frege speaks, instead, in both texts of "true" as a predicate. In the "Logic" of 1897 he writes of it as a predicate which "can be applied to thoughts" (NS, p. 142). And in "The Thought" he writes that "grammatically the word 'true' appears as an adjective" (T, p. 18). Still, the sense-reference account is not altogether forgotten and manifests itself in the characterization of "true" as a predicate applying to "thoughts." Truth, he writes in "The Thought," arises "for the sense of a sentence" or something "I call a thought" (T, pp. 19ff.). The fact that Frege does not call on the doctrine of truth-values at this point is not evidence that he has abandoned it, but rather of a didactic decision to begin with an intuitive account of truth as a property and then to proceed to an exposition of the doctrine of sense and reference. It is this order that he follows also in his notes for Darmstaedter.

Frege's 1897 statement of the indefinability thesis goes together for him with reflection on the notion of fact. He had not employed that notion since the *Begriffsschrift*, where he had suggested that the judgment sign could be read as a predicate, meaning that the judged content "is a fact." The 1897 text reintroduces the notion of fact and links it in a surprising turn to the notion of thought. In Frege's

own words: “Thoughts are, for instance, natural laws, mathematical laws, historical facts” (NS, p. 140). What we think of as facts is thus subsumed under the notion of thought. The same idea recurs in “The Thought,” where he declares apodictically that “a fact is a thought that is true” (T, p. 35). Facts are, on this account, not what thoughts are about—they are themselves thoughts. We tend to speak of thoughts as mental correlates of possible facts; but these would, in Frege’s terminology, be subjective “representations” (*Vorstellungen*) and as such outside the purview of logic. Fregean thoughts are not representations of the world; they rather constitute the world. This may seem odd, at first sight, but is not without its attractions.¹¹ One may plausibly argue that the identity criteria of facts must be intensional, since the fact that Venus is the morning star is surely different from the fact that Venus is the evening star. Facts can therefore, on Frege’s scheme, not be located at the level of reference where identity criteria are unfailingly extensional. It follows from this that truth cannot be conceived as a correspondence between thoughts and facts.

Phase VI: The Primacy of Judgment Reaffirmed (1915)

Frege’s reflections on truth take yet another striking turn in a note from 1915, published posthumously under the title “My Fundamental Logical Insights.” We know nothing about the circumstances under which he wrote it, but its content provides some hints as to Frege’s motivations, for it reverts in a surprising twist to the *Begriffsschrift* conception of logic as primarily concerned with judgment and seeks to reconcile it with the later view of logic as primarily concerned with truth. We may take this to mean that the note was part of an overall stocktaking that Frege undertook as a result of Russell’s discovery of the antinomy in his logical system.

Russell’s communication in 1903 of this discovery had profoundly shaken his earlier assurance that he could ground arithmetic in pure logic. Initially, he seems to have worked on ways to circumvent the antinomy while salvaging the logicist reduction of arithmetic. But as these efforts got nowhere, he appears to have become increasingly despondent. In the text “Logic in Mathematics” from the spring of 1914, the logicist program remains unmentioned. Frege can get himself only to say that “mathematics is more closely related to logic than any other science” (NS, p. 218). And by 1919 he is ready to admit to Darmstaedter the failure of the program. “My Fundamental Logical Insights” was written between these two dates and, thus, at a time of increasing doubt about his achievements. It was natural, then, that Frege should seek to reflect at this moment on his whole development since the *Begriffsschrift*. He characterizes his 1915 note, accordingly, as “a key to the understanding of my results” (NS, p. 271). As such it can be read as the first of his writings from the final period of his life. “The Thought” and the rest of his “Logical Investigations,” his notes for Darmstaedter and the various uncompleted drafts, everything he wrote from now on was an attempt to recuperate what was right in the earlier work, to pull the strands of his thought together, and to show the directions in which they might be extended.

The most striking feature of “My Fundamental Logical Insights” is its already noted reversion to the *Begriffsschrift* conception of logic. When we look at Frege’s writings between 1879 and 1915, we may easily come to think that he had simply

set the idea of logic as concerned with judgment aside in favor of the unqualified belief that its fundamental concern is the notion of truth. But the apparently forgotten thought is unexpectedly reinstated and argued for in 1915, and one is left with the possibility that Frege had never abandoned it completely.

In “My Fundamental Logical Insights” he takes his departure from the observation that a thought is true even before it is grasped by any human being—something he had said many times since 1884. But now he adds to this an attack on the assumption that the word “true” is an adjective in the ordinary sense. He concludes rather that “the word ‘true’ has a sense that contributes nothing to the sense of the whole sentence in which it occurs as a predicate” (NS, p. 272). We know this claim as the “redundance view of truth”—the doctrine that the sentences “A” and “A is true” mean the same. Interpreters often see in it the logical ground for the argument in “The Thought” that truth is indefinable. But “My Fundamental Logical Insights” tells a somewhat different story. For it sets out to justify the redundancy view of truth in terms of another and obviously more basic assumption according to which to *assert* some proposition A means to assert that A is true. Frege writes: “When I assert ‘it is true that sea water is salty’ I assert the same thing as when I assert ‘sea water is salty.’ Through this one recognizes that the assertion does not lie in the word ‘true,’ but in the assertive force with which the sentence is uttered” (NS, p. 271). He comes back to exactly this point three years later when he argues in “The Thought” that any definition of truth must presuppose a practice of assertion. Given any definition of “P is true” as “Q,” we cannot say that P is actually true unless we are already in a position to assert Q; but since the practice of assertion incorporates a grasp of the notion of truth, the attempted definition will turn out to be circular. It presupposes an implicit understanding of that which it tries to define explicitly. When we use the word “true,” we are, according to Frege’s statement in 1915, merely trying “to make what corresponds to the assertive force appear to be a contribution to the thought.” But that is impossible. “And although this attempt fails, or rather because of its failure, it points to the uniqueness of logic” (NS, p. 272). We are thus clearly back in the framework of the *Begriffsschrift*, since the uniqueness of logic of which Frege is speaking here is its fundamental concern with assertion and, hence, with judgment.

But the 1915 text is not simply a reversion to the earlier conception of logic. Frege is now engaged in the challenging project of reconciling that conception with the later idea that truth is primary for logic. He does so in a radically unexpected way by asserting that objectively speaking assertion is, indeed, the primary notion of logic, but that we cannot avoid talking about truth because of the imperfection of our language. We need a notion of truth, in particular, to make the transition from the imperfect language of everyday life to the more perfect language of the *Begriffsschrift*. He writes therefore: “How is it then that this word ‘true,’ though it seems devoid of content, cannot be dispensed with? Would it not be possible, at least in laying the foundations of logic, to avoid this word altogether? That we cannot do so is due to the imperfection of language. If our language were logically more perfect we would perhaps have no further need of logic, or we might read it off from the language” (NS, p. 272).

These are, indeed, radical ideas, but their bearing on Frege’s subsequent work is not immediately obvious. For the first pages of “The Thought” put once again the notion of truth at the center of logic. And the centrality of that notion is em-

phasized even more strongly in the notes to Darmstaedter, where Frege writes: “What is distinctive about my conception of logic is that I give primacy to the content of the word ‘true,’ and then immediately go on to introduce a thought as that to which the question ‘Is it true?’ is in principle applicable” (NS, p. 273). Neither in “The Thought” nor in these notes does Frege refer to the possibility that judgment (or assertion) rather than truth is the primary concern of logic. One might therefore conclude that “My Fundamental Logical Insights” was merely an aside, a dead end, a road not taken. But when one attends more closely to Frege’s insistence in “The Thought” that truth is not a property of propositions, that it is indefinable and unique, one begins to understand that he might still be gesturing in this essay toward the views contained in the note from 1915. According to it, talk about truth is always of a preliminary and didactic nature; any discussion of logic needs to start from that notion but will ultimately have to set it aside. Talk about truth cannot amount to a systematic theorizing; it can only prepare us for the use of a fully worked out logical language, but in such a language the concept of truth has no place. It will manifest itself there only in our ability to make reasoned and justified assertions. On this understanding “My Fundamental Logical Insight” is not the preface to “The Thought” but the coda to which the whole of the “Logical Investigations” lead up.

Phase VII: The Critique of the Picture Theory of Truth (1918)

We can be sure that Frege wrote “The Thought” with Wittgenstein in mind. For (like Russell) he believed him to be the one most likely to continue the work he had begun. While at work on his essay, he wrote to Wittgenstein in April 1918: “What you have gained from our interchanges will hopefully advance humanity a bit on its assigned road. It will be a comforting prospect for me, if the words I have exchanged with you will live on in their effects” (B, p. 16). And he repeated that thought once more in the following year when he foresaw Wittgenstein standing up one day for “what I believe to have discovered in the area of logic.” Their exchange of ideas was important to him, since “in long conversations with you I have got to know a man who like me has searched for truth, though partly on different paths” (B, p. 21).

“The Thought” may be read as Frege’s attempt to show Wittgenstein on what paths he himself has searched for truth. The essay was written and published months before Frege had a chance to set eyes on Wittgenstein’s *Tractatus*, but their correspondence reveals that he expected to find both agreements and disagreements with his own ideas in that text. He knew, for instance, of Wittgenstein’s objections to his doctrine of truth-values from letters they had exchanged in 1913.¹² In conversations Wittgenstein had, moreover, defended Russell’s theory of signs, and that theory included a commitment to a correspondence notion of truth. After having taken truth to be simple and undefinable in the early years of the century, Russell had felt “driven back to *correspondence with fact* as constituting the nature of truth,” as he wrote in 1912 when he added the proviso, however, that “it remains to define what we mean by ‘fact,’ and what is the nature of the correspondence which must subsist between belief and fact, in order that belief may be true.”¹³ This was the project to which Wittgenstein had devoted himself by developing his own “picture-theory of meaning.” The world, he had argued, consists of facts, and a fact

is the existence of a state of affairs which in turn are combinations of objects. The combination of names that makes up a sentence is also a fact, and a sentence is true if it stands in a strict mapping or picturing relation to a fact in the world. The sentence “does not involve a correlation of a fact with an object, but rather the correlation of facts by means of the correlation of their objects.”¹⁴

We cannot say how much of this doctrine Frege actually knew when he wrote “The Thought,” but the essay contains a number of pointedly anti-Tractarian statements. It begins with a critique of the correspondence conception of truth in the form of a picture theory of the sort we find in Wittgenstein’s treatise. As already quoted, Frege argues that the truth of a proposition cannot consist in its pictorial correspondence to reality, since such a correspondence “can only be perfect if the corresponding things coincide” (T, p. 18). But this is neither expected nor desired in the case of propositions. If we say, on the other hand, that the proposition is true when it corresponds to a fact in a certain respect, we will be forced to ask whether it is true that it corresponds in the laid-down respect, and then we presuppose the notion of truth we are seeking to define. These are, indeed, issues that Wittgenstein is forced to address in the *Tractatus*. He writes accordingly that “in the picture and the pictured there must be something identical in order that the one can be a picture of the other at all” (TLP, 2.161). And he postulates a notion of logical form that can be attributed to both propositions and facts, to language and the world. “What every picture, of whatever form, must have in common with reality in order to be able to represent it at all—rightly or falsely—is the logical form, that is, the form of reality” (TLP, 2.18). This is not, however, a notion Frege would have been comfortable with, and his refusal to define truth may be considered an expression of his resistance to it.

Though Frege’s attack on a picture conception of truth may be taken to be directed against the *Tractatus* view, one must not overlook that he had engaged in similar criticisms already in his “Logik” of 1897. He had insisted then that the sentence is a vehicle for the expression of thoughts, not of ideas (*Vorstellungen*), since it is “not well suited to reproduce ideas. Pictures and musical pieces are, by contrast, unsuited to express thoughts” (NS, p. 137). And he had gone on to characterize talk of the truth of ideas and pictures as merely derivative: “An idea, like any other picture, is not true in itself but only in respect to something to which it is supposed to conform. . . . Without reference to an intention to depict something, one cannot talk of the truth of a picture. From this one can derive that it is not properly speaking the idea to which the predicate ‘true’ is attributed, but the thought that it depicts a certain object” (NS, p. 142). Though the conclusion is the same as the one he comes to in 1918, he reaches it here by a different route. And this difference may be explained by the fact in “The Thought” Frege is concerned with the kind of picture conception of truth we find in the *Tractatus*.

When Frege received the actual text of the *Tractatus* at the end of 1918, his reaction to it was decidedly cool. Despite repeated entreaties from Wittgenstein (WBW, p. 267; B, p. 19), he did not respond to it till the following June when he wrote apologetically: “You must certainly have been awaiting an answer from me for some time now and must have wished for a comment from me on your treatise.” To excuse himself he mentioned “complex business matters” and the fact that “I often feel tired now” (B, pp. 19, 20). While he tried to explain thus his inability to give a reasoned judgment on the *Tractatus*, he could not avoid the even-

tual admission: "I find it difficult to understand" (B, p. 19). This was a complaint he would repeat again somewhat later, in August, when he wrote: "I cannot give an assessment of the treatise itself, not because I disagree with its content, but because that content is insufficiently clear to me" (B, p. 23).

Frege was puzzled by the "dogmatic" style of the *Tractatus*, by the omission of arguments and explanations, by the lack of a clearly stated problem to which the text addresses itself, by Wittgenstein's insistence that the reader must already have entertained the thought expressed in the work, and by Wittgenstein's aesthetic intentions. "What you write to me about the purpose of the book, is strange to me," he remarked in September 1919. "I proceeded in my comments from the assumption that you wanted to communicate a new content. And in that case the greatest clarity would be the greatest beauty" (B, p. 21). It was clear from all this that they would never come to agree with each other. While Frege's last letter to Wittgenstein, from April 1920, appealed once more to their "old friendship," they were, in fact, going their separate ways.

At the center of Frege's critique of both Russell's correspondence theory and Wittgenstein's picture theory of truth stands an attack on the notion of fact the two had employed. The observation in "The Thought" that a fact is a thought that is true, far from being a casual aside, is for that reason of critical importance for Frege. He makes this evident in the criticisms of the *Tractatus* he eventually offers Wittgenstein in his letters. Interpreters have occasionally downplayed these comments because Frege prefaces them with an admission that he finds Wittgenstein's treatise difficult to understand. But this is surely meant to be taken in a philosophical rather than a pedestrian sense. Frege concentrates in his critique on the first pages of the *Tractatus*, writing to Wittgenstein: "You use right at the beginning a fairly large number of words on whose precise sense obviously much depends" (B, p. 19), and he singles out terms like "being the case," "state of affairs," "situation," and above all "fact" as being problematic. He expresses puzzlement, in particular, at Wittgenstein's statement that a fact is the existence of states of affairs. "I understand this to mean that every fact is the existence of a state of affairs, such that another fact is the existence of another state of affairs. Would it be possible to strike out the word 'existence' and say: 'Every fact is a state of affairs, every other fact is another state of affairs?'" (B, p. 20). These are concerns that later readers of the *Tractatus* have also felt.

Frege continues that these initial doubts have made it difficult for him to advance in his reading of the treatise. His predicament is, however, not merely terminological in nature, as the further course of his argumentation reveals. Since Wittgenstein characterizes states of affairs as combinations of objects, Frege asks whether that means that facts, too, are combinations of things. This is important to him because it implies that facts have extensional identity criteria. He refers in this connection to the principle that the part of a part is a part of the whole, and so, if facts are composed of objects, they must also be composed of the parts of these objects. He writes: "I would like to have an example for the Vesuvius being a component of a state of affairs. That would mean, it seems to me, that the components of the Vesuvius are components of this fact; the fact will thus consist of hardened lava. That does not seem to me right" (B, p. 20). His argument comes to this: the state of affairs that the Vesuvius is near Naples is clearly different from the state of affairs that a heap of hardened lava is near Naples, for the first may be

the case even if the Vesuvius does not consist of hardened lava. The identity criteria for states of affairs and facts must therefore be intensional. But this gets us back to the conclusion that facts are simply true thoughts, and that truth cannot therefore be defined as a relation between thoughts (or the sentences that express thoughts) and facts.¹⁵

Wittgenstein failed to appreciate these critical comments, just as he failed to appreciate the alternative view Frege had laid out in "The Thought." Wittgenstein had received a copy of that essay after his release from the prisoner-of-war camp on August 21, 1919. But he found it just as difficult to respond to it as Frege had found it to assess the *Tractatus*. His life, too, was in turmoil, and so in September 1919 Frege, in turn, found it necessary to nudge him gently for a response by pointing out convergences between Wittgenstein's views and his own. Only in March 1920 did Wittgenstein finally get around to addressing Frege's concerns, and then in such a highly critical manner that he felt it necessary to apologize in advance. "Naturally, I do not resent your outspokenness," Frege replied. But he was utterly puzzled by Wittgenstein's comments and, obviously suspicious that Wittgenstein had not seriously studied his essay, urged him "to go through the article on 'The Thought' to the first sentence with which you disagree and to write it down together with the reasons for your deviation." (B, p. 24). It looks as if Wittgenstein's exchanges with Frege had come to an end at this point. While he retained a general admiration for Frege, he never seems to have come around to appreciating "The Thought" because in later years he dissuaded Max Black and Peter Geach from including it in the collection of Frege's writings that he encouraged them to publish.

Despite their patent disagreements, there was, however, one fundamental point on which Frege and Wittgenstein might have come to a mutual understanding at this point, for they both entertained similar doubts about the possibility of a *theory* of truth. Frege had expressed his doubt most directly in "My Fundamental Logical Insights," which Wittgenstein did, of course, not know, since it remained unpublished until 1969. If he had been acquainted with that text, he might have been surprised by the almost Tractarian tone of Frege's conclusion that in a fully functioning language talk about truth would eventually drop away. If Frege had not been put off by the first pages of the *Tractatus*, he, in turn, might have come to appreciate how close the final pages of that work come to his 1915 arguments. While the early parts of the *Tractatus* spell out an account of truth, Wittgenstein proceeds there to deconstruct all semantic theorizing and concludes that all attempts to speak *about* logic are bound to fail. The result is the concluding declaration in the *Tractatus* that "my propositions serve as elucidations in the following way: anyone who understands me eventually recognizes them as nonsensical. . . . He must transcend these propositions, and then he will see the world aright" (TLP, 6.54). For on Wittgenstein's conception, "logic is not a theory but a reflexion of the world," and as such "transcendental" (TLP, 6.13, and 6.2). Similarly to the Frege of 1915, he concludes in a memorable phrase that "logic must take care of itself" (TLP, 5.4731).

III. Frege against Tarski

Thirty years ago, Jean van Heijenoort argued in a seminal article that none of the original authorities in the rise of mathematical logic, neither Frege nor Russell

nor Wittgenstein, had been concerned with questions of semantics, whose formal development had been left to another generation. The “logicists” had taken logic instead as a universal language such that “nothing can be, or has to be, said outside the system.” In consequence, Frege had never raised “any metasystematic question,” and “questions about the system are as absent from *Principia Mathematica* as they are from Frege’s work. Semantic notions are unknown.”¹⁶

It is certainly true that none of the logicists constructed a *formalized* metalanguage or a *formalized* theory of meaning or truth. What we have in their writings are incidental arguments that are sometimes disowned as soon as they are made. But van Heijenoort overstated his case by ignoring Frege’s, Russell’s, and Wittgenstein’s extensive discussions of the notions of meaning and truth. There is, in fact, no other period in the whole history of logic in which the concept of truth was examined with such passion and in which such different proposals emerged about how it should be understood. And this was no accident. Since the old Aristotelian logic had been tied to the classical correspondence conception of truth, the logicists found it necessary to rethink, revise, or reject the traditional conception before they could make advances in their own logical investigations.¹⁷

It has been said that Tarski’s definition of truth returns us to the prelogicist tradition, that it salvaged the Aristotelian notion of truth and showed how it could be made precise, that Tarski achieved what the logicists either neglected or claimed to be impossible. This would imply that he opened the door once again for the kind of metaphysical speculation that surrounded the classical notion of truth. But there is another and more compelling way to read Tarski. It holds that his truth-definition has nothing to do with the traditional conception of truth as correspondence to the world. Tarski offers us, instead, an account of how sentences of one language, the object-language, are correlated with those of another, the metalanguage. On this account Tarski’s criterion of adequacy declares in metalinguistic terms that an object-language sentence can be called true when we are in a position to assert its metalinguistic translation. No metaphysical consequences follow from this.

Understood in this minimalist fashion, Tarski’s theory may be taken to confirm the Fregean intuition that assertion is primary to logic rather than truth. However, if the notion of truth is contained in that of assertion, it also holds for both Tarski and Frege that we must understand the notion implicitly in one language (the metalanguage) before we can explain it explicitly in another. Tarski can offer us therefore at most a partial definition of the concept of truth, and since Frege opposed any kind of partial definition, he would have had to reject also Tarski’s formal construction. There are additional reasons why Frege would not have been happy with Tarski’s theory. His claim that words have meaning only in the context of a sentence and his related assertion of a categorial difference between functions and objects rule out any theoretical account of the truth of sentences in terms of the meanings of their components. That these doctrines are not reconcilable with Tarski’s theory can be seen from Rudolf Carnap’s initial resistance to formal semantics, which stemmed from his early reliance on Fregean assumptions. Carnap changed his mind on the prospects of a Tarski-style semantics only after he had abandoned these inherited commitments.¹⁸

Most logicians today will think that Carnap did right in this. But even if we concede this point and allow for the legitimacy of Tarski’s technical constructions,

we are forced to conclude that there remain insights to be gleaned from Frege's alternative line of reasoning. One of these is the observation that the concept of truth is so fundamental to our understanding that it is impossible to reduce it to other, more primitive notions. A second concerns the normative character of the notion of truth, a topic on which Tarski's theory has nothing to say. While we may want to resist Frege's own value-theoretical explanation of this normativity, we still need to account for it in some way or another. Finally, there is the insight which Frege shared with Wittgenstein, that our practice of assertion is certainly responsible to the world, but that the attempt to explicate the relation between language and world must remain problematic because it is, of necessity, always confined within the bounds of language.

Notes

I have learned much from questions and comments on earlier versions of this essay. I am particularly grateful to Jamie Tappenden for extensive notes on a somewhat earlier draft. I wish that time and space had allowed me to address the serious concerns and views he has expressed in Tappenden, 1997, in the manner they deserve.

1. Frege, 1967b, p. 19 (hereafter referred to as "T"). In this passage Frege characterizes his conclusion as merely "probable," but this is not to be taken as a sign of his uncertainty about it. He is expressing rather his realization that it depends on a long string of considerations that go all the way back to the *Begriffsschrift* and not simply on the sketchy argument he provides in "The Thought." This is, in any case, the assumption on which I proceed in this discussion.

2. Frege, 1989, p. 18 (hereafter referred to as "B").

3. On Bauch, his journal, and the Deutsche Philosophische Gesellschaft, see Sluga, 1993.

4. Frege, 1967a, p. 5 (hereafter referred to as "BS").

5. See Sluga, 1987.

6. Transcribed from Frege, 1969, p. 47 (*Nachgelassene Schriften* hereafter referred to as "NS").

7. Frege, 1959, pp. 2–4 (hereafter referred to as "F").

8. Frege, 1972, p. 84.

9. Windelband, 1882, p. 3; see also Windelband, 1884. The observation that the term "truth-value" is due to Windelband was first made by Gottfried Gabriel in Gabriel, 1986. For a discussion of these issues, see also Sluga, 1996. Given Frege's interest in the nature of negative judgments, it seems to me likely that he discovered the notion in Windelband's 1884 essay rather than in the less accessible earlier piece. Windelband writes in 1884 (in an admittedly psychologistic language) of the need to distinguish in every judgment between the mere combination of ideas and an act of adjudication (*Beurtheilung*) that concerns "the truth-value of the judgment" (p. 170), adding that "the truth-value must be correlated with the other values" (p. 174).

10. Rudolf Carnap recorded an even more abbreviated version of the argument when he attended Frege's lectures at Jena in the winter of 1910. Frege concluded then that "truth cannot be defined, analyzed, or reduced. It is something simple, primordial" (Frege, 1996, p. 15).

11. Frege's language is certainly no more odd than G. E. Moore's doctrine in his seminal 1899 essay "The Nature of Judgment" that the world consists of judgments or Russell's related claim in *The Principles of Mathematics* that it consists of propositions; see Moore, 1899; Russell, 1959a.

12. Frege, 1976, p. 266 (hereafter referred to as "WBW".)

13. Russell, 1959, p. 123.

14. Wittgenstein, 1922, 5.542 (hereafter referred to as "TLP").

15. There are some other notable disagreements between "The Thought" and the *Tractatus* that cannot be discussed here. In contrast to Wittgenstein but in agreement with Russell, Frege maintains in "The Thought" that there are real subjects in the world and that each subject has an incommunicable knowledge of itself. And in further contrast to Wittgenstein, but once again in agreement with Russell, he argues that logical truths have substantive content and are not mere tautologies.

16. van Heijenoort, 1993, p. 74.

17. Sluga, forthcoming.

18. Sluga, 1999.

Bibliography

- Frege, G. (1959). *Die Grundlagen der Arithmetik*, with a translation by J. L. Austin. Oxford: Blackwell.
- . (1967a). *Begriffsschrift*. In *From Frege to Gödel*, J. V. Heijenoort, ed. Cambridge, Mass.: Harvard University Press, pp. 1–82.
- . (1967b). "The Thought." In *Philosophical Logic*, P. F. Strawson, ed. M. Quinton, trans. Oxford: Oxford University Press, pp. 17–38.
- . (1969). "Booles rechnende Logik und die Begriffsschrift." In *Nachgelassene Schriften*, H. Hermes et al., eds. Hamburg: Meiner, pp. 9–52.
- . (1972). "On the Scientific Justification of a Begriffsschrift." In *Conceptual Notation*, T. W. Bynum, ed. and trans. Oxford: Oxford University Press, pp. 83–89.
- . (1976). *Wissenschaftlicher Briefwechsel*. G. Gabriel et al., eds. Hamburg: Meiner.
- . (1989). "Briefe an Ludwig Wittgenstein." A. Janik, ed. in *Grazer Philosophische Studien* 33/34:5–34.
- . (1996). "Vorlesungen über Begriffsschrift." G. Gabriel et al., eds. *History and Philosophy of Logic* 17:1–48.
- Gabriel, G. (1986). "Frege als Neukantianer." *Kantstudien* 77:84–101.
- Moore, G. E. (1899). "The Nature of Judgment." In *G. E. Moore: Selected Writings*, T. Baldwin, ed. London: Routledge, pp. 176–93.
- Russell, B. (1959a). *The Principles of Mathematics* [Originally London: Allen and Unwin, 1903].
- . (1959b). *The Problems of Philosophy*. Oxford: Oxford University Press.
- Sluga, H. (1987). "Frege against the Booleans." *Notre Dame Journal of Formal Logic* 28:80–98.
- . (1993). *Heidegger's Crisis: Philosophy and Politics in Nazi Germany*. Cambridge, Mass.: Harvard University Press.
- . (1996). "Frege on Meaning." *Ratio* 9:209–26.
- . (1999). "Truth before Tarski." In *Alfred Tarski and the Vienna Circle*, J. Wolenski and E. Köhler, eds. Dordrecht: Kluwer.
- . (forthcoming). "Truth from Bolzano to Frege." Festschrift für Kurt Rudolf Fischer, Vienna.
- Tappenden, J. (1997). "Metatheory and Mathematical Practice." *Philosophical Topics* 25:213–64.
- van Heijenoort, J. (1993). "Logic as Calculus and Logic as Language." In *The Philosophy of Frege*. Vol. 1. H. Sluga, ed. New York: Garland, pp. 72–78.
- Windelband, W. (1882). "Was ist Philosophie?" In *Präudien*. Vol. 1. Freiburg: Mohr, 1921.
- . (1884). "Beiträge zur Theorie der negativen Urteile." In *Strassburger Abhandlungen zur Philosophie*. Freiburg: Mohr.
- Wittgenstein, L. (1922). *Tractatus Logico-Philosophicus*. C. K. Ogden, trans. London: Routledge and Kegan Paul.