

CHAPTER TWO

TRUTH II

The Twentieth Century, Necessity, and Possible Worlds

The Twentieth Century

BEGINNING IN THE EARLY TWENTIETH CENTURY, TWO NEW CONCEPTIONS of truth arose. These views held that the long-standing traditionalist project of identifying the nature of truth was at best in need of revision and at worst should be abandoned. The traditionalist search for the deep essence or nature of truth was misguided. **Deflationism** holds that there is no special property or nature of truth for a theory of truth to reveal. Identifying some belief or sentence as “true” is redundant or a mere linguistic convenience. Several prominent twentieth-century philosophers have held some version of the view, and versions of the view are held by many philosophers today. At least one person has remarked that some version of deflationism is now the most widely held view of truth (rather than some version of correspondence).

Much of twentieth-century philosophy was motivated by both a generally empiricist outlook and a preference for a more sparse metaphysics, one that invoked fewer types of objects and properties. Moreover, as noted in the previous chapter, a recognition of the importance of language, specifically our understanding of the nature of linguistic meaning, shaped theories across a wide range of philosophical disciplines. Truth theories were not immune. The idea that *truth* or *being true* names a special kind of property was suspect. As philosophers developed theories of truth that shunned a special “truth property,” deflationism became more widely endorsed and accepted.

One influence on the deflationist approach stems from Wittgenstein, who proposed a very different view of the task of philosophy, in his two most influential works, the earlier *Tractatus* (see Chap. 1) and the later *Philosophical Investigations*. In both works, Wittgenstein is suspicious of the idea that philosophy reveals to us the real nature of things, including truth, propositions, meaning. Many regard both works as anti-metaphysical. The important point for our purpose is that in his view we can be misled by language (although he takes different tacks in the two works). In the *Investigations*, he remarks that “philosophical problems arise when language goes on holiday.”¹ Later he claims “Philosophy is a battle against the bewitchment of our intelligence by means of language.”² Wittgenstein thought that the structure of language leads to confusion: “One thinks one is tracing the outline of the thing’s nature over and over, and one is merely tracing round the frame through which we look at it.”³ His aim is to “shew the fly the way out of the fly-bottle” by reminding us that philosophy “leaves everything as it is” and is not uncovering some hidden nature. We think there is a deep essence to concepts such as meaning or truth. Instead, we need, according to Wittgenstein, to remind ourselves of the many different ways language can be used and hence the ways in which we might be “bewitched.” A simple example (though not one particularly relevant to philosophy) of how language can “bewitch”: consider *sakes*. What is a sake, really? We think that there must be one essence common to its different uses: “for the sake of your health,” “for your children’s sake,” “for God’s sake.” It seems to be something that can be had by other people and other things, but what exactly is it? How many of them are there? Could there be sakes without people? (If this reminds you a bit of the disappearing Cheshire Cat leaving

1 1953, p. 38.

2 *Ibid.*, p. 109.

3 *Ibid.*, pp. 309, 114.

behind only a smile, you might be interested to know that Wittgenstein was very familiar with some of Lewis Carroll's work.) But these are silly questions that shouldn't have been asked in the first place, if we had only attended to the manifest ways that we use that phrase in ordinary language and were not misled by the surface grammar. The search for "essence" or "natures" is of course held by some to be the defining *nature* of metaphysics. But some deflationists about truth think that questions about the nature of truth arise from a similar sort of "bewitchment." Whatever their motivation, deflationists hold that there is no special property or characteristic property that is the essence of truth.

A hugely important contribution, not only to the topic of truth, but also to other areas of philosophy, appeared in 1935:⁴ Alfred Tarski's article "Der Wahrheitsbegriff in Den Formalisierten Sprachen," later translated into English as the "The Concept of Truth in Formalized Languages." In the essay, Tarski's principal concern was to understand the notion of truth as it is used in the language of mathematics (and logic). In 1944, he published a second essay on the same topic, "The Semantic Conception of Truth."⁵ It is this essay that gives us the name of this view of truth. The **semantic conception of truth** sees truth as essentially tied to a language and relies on the notions of *object language*, *metalinguage*, and *satisfaction* (explained below) to elucidate what it is to be a true sentence. Both of these essays *formalize* the notion of truth, using an advanced logic apparatus.

The semantic conception is not the first version of deflationism, but is an appropriate place to begin consideration of that general movement. Versions of deflationism rely on an important feature of the semantic conception.

Semantic Conception of Truth

The development of formal logic in the late nineteenth and early twentieth centuries provided very powerful tools of analysis and theorizing to philosophers, mathematicians, and logicians. For example, in 1905 Bertrand Russell made use of predicate logic in order to demonstrate a difference in the logical structure of sentences containing names and sentences containing definite descriptions (descriptions that purport to denote a unique individual, e.g., 'the present president of the US'). Mathematicians and logicians are of course interested in notions such as truth and consistency. Prior to Tarski's work,

4 Tarski 1935.

5 Tarski 1944.

there was no rigorously formulated account of truth for formal systems, that is, mathematical or logical languages. When Tarski set out to show how truth could be “formalized,” he wanted to do so without relying on any semantic notion, such as truth or reference. But he also had in mind apparent paradoxes about truth, including the ancient and notorious “Paradox of the liar.” How should we understand someone who says “I am lying” or an instance of the sentence “This sentence is false”? (The sentence seems to refer to itself, saying that it’s false; yet that’s just what the sentence says. So, it’s true. But if it’s true, then isn’t it false?)

Without entering into the formal apparatus employed by Tarski, we should be able to get a pretty good sense of his notion of “truth-in-a-language.” First point: any sentence is true *relative to a language*. In one sense, this is to be expected. Remember, Tarski was working with formal languages, where the domain—what the language is about—can be specified with some precision. Then, for example, “geometry sentences” are true (or false) relative to the language we are using to talk about geometric objects. Some think this is a disadvantage of the semantic conception: it does not give us an account of truth in general. And it is sometimes asked whether the semantic conception of truth can be applied to *natural languages*, languages such as English or Spanish or Italian or Polish. As we will see later, a prominent account of the meaning of natural language sentences utilizes the Tarskian idea.

What is it to be true for a formal language sentence? (We will be rather informal here, since it is the general Tarskian thought that we want.) Start with a domain—the objects or things to be talked about. In arithmetic, we talk about doing things with numbers; in geometry, it’s geometric objects. And our “framework” for talking about those objects is a way to name them, both generally and specifically. The axioms, together with some basic logic rules, tell us the sentences we can generate. Now, select some sentence. Let’s try the Pythagorean Theorem: “For every right triangle, the square of the hypotenuse is equal to the sum of the squares of the two adjacent sides.” This sentence is true if it is *satisfied by* all the right triangle objects in the domain. Why “all”? Because this is a *universally quantified* sentence; it’s about all the right triangles. And this places a constraint on the sentence being satisfied. Thus, the Pythagorean Theorem is true if and only if all the right triangles meet the condition set out by the theorem. Tarski tells us (very metaphorically): go find all the right triangles; do they *satisfy* the sentence? Do they meet the condition set out by the theorem?



One more example from arithmetic: There is a successor (a number that follows) of 7 that is less than 10. Notice this is not a universal sentence; it's a "some" sentence, or an "existentially quantified" sentence. It says "There is at least one..." Is it satisfied? Well (metaphorically again): go look at the numbers. Is there at least one number that follows 7 but is less than 10? Yes, there is. Thus, the sentence is satisfied. (Tarski's own formulation was more complicated, but we need not worry about that here.)

We now have an account of truth for sentences in a formal language. Universal sentences are "true-in-L" (the formal language we are using) if and only if all relevant objects in the domain **satisfy** the sentence (i.e., their insertion makes the sentence true). Can't get complete satisfaction? It's false. For a "some" sentence—an existentially quantified sentence—it is "true-in-L" if and only if at least one object in the domain satisfies it. Can't get no satisfaction? It's false. "True-in-L" is satisfaction.

Tarski wanted to explain truth without relying on the concept of truth itself or some concept that presupposed the concept of truth. And many think he succeeded in doing so. Some, however, suspect that satisfaction is itself a disguised semantic notion. And Tarski seems to suggest that his account

of truth is a kind of correspondence theory; others have similarly claimed that Tarski's semantic approach is a correspondence theory. However, there is no attempt to offer an explanation or definition of "correspondence." Tarski viewed his account as a "material adequacy" condition. Any satisfactory or "adequate" theory of truth should yield the sort of connection offered by T-sentences.

Could you apply this approach to truth in natural languages? Perhaps we can get a sense of what that might be like for a restricted domain for English, say, the domain of *human beings*. As a universal sentence, consider the sentence S_1 : "All human beings are mortal." Does every object in the domain—every human being—satisfy the condition? The answer is "yes." Then this sentence is "true-in-English." Or S_2 : "Someone knows where Alessia is." Is there at least one object in the domain that satisfies the sentence? Again, if the answer is "yes," then the sentence is true. One more: "Tarski taught at the University of California, Berkeley." This sentence is satisfied if the object in the domain that is picked out by the name "Tarski" has the property or characteristic identified by the predicate "taught at the University of California, Berkeley." He did. So the object named by "Tarski" has the requisite property, and thus the sentence is true. Again, this sentence is *true-in-English*.

Notice that this latter sentence seems to fit with the idea of correspondence. The sentence is true because it is satisfied, which someone might think is a different way of talking about correspondence. But others see coherence at work; after all, it is truth relative to the *whole* language. A sentence is true relative to the framework of the language as a whole. We leave this issue aside.

One further important point. Tarski showed—and required—that the truth conditions of any given sentence in a formal language be derived from a small set of axioms together with a set of names of the objects and a set of primitive predicates, or the names for the primitive properties. Thus, there is a theorem for each and every sentence in the language that spells out that sentence's *truth conditions*. These derivable sentences—theorems—spell out the truth conditions in the form of biconditionals, or if and only if sentences.

Suppose for a moment you could axiomatize English and provide the required list of names (and variable names) and primitive predicates. A Tarskian account of "true-in-English" would provide a derivation, a theorem of this form:

T(S): "S" is true-in-English if and only if...

where “S” is the name of a sentence, and the ellipsis is filled in by the truth conditions of S, whatever those happen to be. These are called the T-sentences of English.⁶ Remember: if Tarski’s approach works, you could derive a T-sentence for each English sentence. And “true-in-English” is then nothing more than just this very long (at least countably infinite) list of T-sentences or theorems.

(“Countably infinite” groups or sets can be counted using the natural numbers: 1, 2, 3,.... You will never run out of numbers, but it is an unending task. And, yes, there are “uncountably infinite groups or sets,” for example, the real numbers, as the nineteenth-century mathematician, Georg Cantor, proved.)

Perhaps the most frequently cited T-sentence, “Snow is white,” is true-in-English if and only if snow is white. This particular T-sentence looks trivial and uninformative. Yet the *derivation* of this sentence is far from trivial (depending on the axioms). Moreover, consider a T-sentence for the French sentence “La neige est blanche.” That is:

“La neige est blanche” is true-in-French if and only if snow is white.

This leads us to see the difference between an **object language**, the language we are attempting to specify the range of T-sentences, e.g., the language of arithmetic, and the **metalanguage**, the language we are using to talk about the object language. Here we have French as the object language and English as the metalanguage. In a formal language, such as that of arithmetic, the object language is a restricted fragment of English, together with the required mathematical concepts.

Tarski noticed that natural languages like English are at once object language and metalanguage. When I tell my friend, “In English, ‘see you later’ means ...,” I am using English as the metalanguage to say something about English, the object language. My explanation is a level above the simple “see you later.”⁷

6 See Richard Montague 1974 for examples of the attempt to axiomatize fragments of English, using modal logic.

7 And it is this distinction between object and metalanguage that allows Tarski to give an explanation of the Liar Paradox. Tarski’s idea is that to prevent the paradox from arising, one needs to “step outside” the language into a metalanguage to talk about when a sentence is true (or false). This is Tarski’s way of avoiding the self-referential problem in the Liar Paradox. See Bradley Dowden’s “Liar Paradox” for a helpful and detailed account of the paradox.

T-sentences, and their derivability, as noted earlier, turn out to be important for deflationism and have been famously employed by Donald Davidson and advocates of a truth conditional approach to the theory of meaning.⁸ Similarly the notions of meta- and object language figure in arguments that we will see about realism. It is interesting to note that this seemingly innocuous idea, exemplified by “‘Snow is white’ is true if and only if snow is white,” has been utilized by proponents of every major theory of truth, and some of those proponents claim that it lends support to their preferred approach. Tarski’s idea exemplifies something essential about truth. How this idea should be utilized in a theory of truth is another question. In fact, deflationists argue that it shows us that there is not some deep structure to truth, which requires a theory to reveal.

Deflationism

Unlike traditional accounts of truth, deflationism holds that there is no special property or nature of truth to be revealed by a theory of truth. Reference to the nature of truth is meant to convey the thought that there is something that all true sentences have in common, *other than simply being among all the true sentences*.⁹ We shall sketch three different deflationist views: redundancy, disquotationalism, and the minimalist theory.

The motivational thought behind these views is that there is nothing *substantial* to truth, and hence they are sometimes known as “insubstantialists.” However they are termed, they agree that the traditionalists’ search for the deep structure of truth—correspondence, coherence, even pragmatist—is a mistake. The redundancy view is typically considered to be the first deflationary theory to appear.

Redundancy

Announcing that some sentence “is true” can, on at least some occasions, look like it’s merely a “repetition device.” Samantha announces, for example,

8 Davidson 1967.

9 There are “classification issues” that arise when considering deflationism. For example, the minimalist theory of truth, as developed and advocated by Paul Horwich, is sometimes counted as deflationist and sometimes not. For our purposes, the minimalist view is treated here as deflationist, recognizing that Horwich counts “is true” as naming a property. Cp. Frederick Schmitt 1995, pp. 125 ff, and Richard Kirkham, pp. 339 ff.