

CHAPTER 1

## The Need for New Foundations



In this book, I will argue that the revolution in the study of language, mind, and meaning led by advances in philosophical logic from Frege through Tarski, Kripke, Montague, and Kaplan must be reconceptualized. Although much progress has been made by adapting intensional logic to the study of natural language, the resulting theoretical framework has limitations that require rethinking much of what has guided us up to now. I will begin by sketching where we are in the study of linguistic meaning and how we got there, after which I will identify three main ways in which I believe the current theoretical framework must change.

The story begins with the development of symbolic logic by Gottlob Frege and Bertrand Russell at the end of the nineteenth and beginning of the twentieth centuries. Initially, their goal was to answer two questions in the philosophy of mathematics: *What is the source of mathematical knowledge?* and *What are numbers?* They answered (roughly) that *logic* is the source of mathematical knowledge, that *zero* is the set of concepts true of nothing, that *one* is the set of concepts true something, and only that thing, that *two* is the set of concepts true of some distinct  $x$  and  $y$ , and nothing else, and so on. Since the concept *being non-self-identical* is true of nothing, it is a member of zero; since the concept *being the Hempel lecturer in 2013* is true of me and only me, it is a member of the number one;

since the concept *being my son* is true of Greg and Brian Soames, and only them, it is a member of the number two. Other integers follow in train. Since numbers are sets of concepts, the successor of a number  $n$  is the set of concepts  $F$  such that for some  $x$  of which  $F$  is true, the concept *being an  $F$  which is not identical to  $x$*  is a member of  $n$ . Natural numbers are defined as those things that are members of every set that contains zero, and that, whenever it contains something, always contains its successor. Multiplication is defined as repeated addition, while addition is defined as repeated application of the successor function. In this way arithmetic was derived from what Frege and Russell took to be pure logic. When, in similar fashion, classical results of higher mathematics were derived from arithmetic, it was thought that all classical mathematics could be so generated. So, logic was seen the foundation of all mathematical knowledge.

That, at any rate, was the breathtaking dream of Frege and Russell. The reality was more complicated. Their first step was the development of the *predicate calculus* (of first and higher orders), which combined truth-functional logic, familiar from the Stoics onward, with a powerful new account of generality supplanting the more limited syllogistic logic dating back to Aristotle. The key move was to trade the subject/predicate distinction of syllogistic logic for an expanded version of the function/argument distinction from mathematics. Applied to quantification, this meant treating the claim *that something is  $F$*  as predicating *being true of something* of the property *being  $F$*  or, in Russell's convenient formulation, of the function that maps an object onto *the proposition that it is  $F$* , while treating the claim *that everything is  $F$*  as predicating *being true of each object* of that property or function. The crucial point, resulting in a vast increase in expressive power, is the analysis of *all* and *some* as

expressing higher-order properties of properties or propositional functions expressed by formulas of arbitrary complexity.<sup>1</sup>

Although the first-order fragment of Frege's system was sound and complete—in the sense of proving all and only genuine logical truths—the concepts needed to define and prove this (while also proving that the higher-order system was sound but, like all such systems, incomplete) were still fifty years away. In itself, this didn't defeat the reduction of mathematics to logic. More serious was the intertwining of this early stage of modern logic with what we now call “naïve set theory”—according to which for every stateable condition on objects there is a set (perhaps empty, perhaps not) of all and only the things satisfying it. To think of this as a principle of logic is to think that talk of something's *being so-and-so* is interchangeable with talk of its being in the set of so-and-so's.

When Russell's paradox demonstrated the contradiction at the heart of this system, it quickly became clear that the principles required to generate sets without falling into contradiction are less obvious, and open to greater doubt, than the arithmetical principles that Frege and Russell hoped to derive from them. This undercut the initial epistemological motivation for reducing mathematics to logic. Partly for this reason, the subsequent boundary that grew up between logic and set theory was one in which the latter came to be viewed as itself an elementary mathematical theory, rather than a part of logic. Reductions of mathematical theories to set theory could still be done, with illuminating results for the foundations of mathematics, but the philosophical payoff was not what Frege and Russell initially hoped for.<sup>2</sup>

<sup>1</sup> 'F' and 'G' are here used as schematic letters.

<sup>2</sup> This story is told in much greater detail in chapters 1, 2, 7, and 10 of Soames (2014b).

This *philosophical* shortcoming was compensated by the birth of new deductive disciplines—proof theory and model theory—to study the powerful new logical systems that had been developed. A modern system of logic consists of a formally defined language, plus a proof procedure, often in the form of a set of axioms and rules of inference. A proof is a finite sequence of lines each of which is an axiom or a formula obtainable from earlier lines by inference rules. Whether or not something counts as a proof is decidable merely by inspecting the formula on each line, and determining whether it is an axiom, and, if it isn't, whether it bears the structural relation to earlier lines required by the rules. Since these are trivially decidable questions, it can always be decided whether something counts as a proof, thus forestalling the need to prove that something is a proof. In a purely logical (first-order) system, the aim is to prove all and only the *logical truths*, and to be able to derive from any statement all and only its *logical consequences*.

These notions are defined *semantically*. To think of them in this way is to think of them as having something to do with *meaning*. Although this wasn't exactly how the founder of model theory, Alfred Tarski, initially conceived them, it is how his work was interpreted by Rudolf Carnap and many who followed. The key idea is that *we can study the meaning of sentences by studying what would make them true*. This is done by constructing abstract models of the world and checking to see which sentences are true in which models. When a sentence is true in all models it is a logical truth; when the truth of one sentence in a model always guarantees the truth of another, the second is a *logical consequence* of the first; when two sentences are always true together or false together they are *logically equivalent*, which is the logician's approximation of sameness of meaning.

By the mid-1930s, the model and proof theories of the first- and second-order predicate calculi were well understood and inspiring

new projects. One was modal logic, which introduced an operator *it is logically/analytically/necessarily true that*—the prefixing of which to a standard logical truth produces a truth. Apart from confusion about what logical, semantic, or metaphysical notion was to be captured, the technical ideas soon emerged. Since the new operators are defined in terms of truth at *model-like elements*, logical models for modal languages had to contain such elements, now dubbed *possible world-states*, thought of as *ways the world could have been*. This development strengthened the Fregean idea that for a (declarative) sentence *S* to be meaningful is for *S* to represent the world as being a certain way, which is to impose conditions the world must satisfy if *S* is to be true.

Hence, it was thought, meaning could be studied by using the syntactic structure of sentences plus the representational contents of their parts to specify their truth conditions. With the advent of modality, these conditions were for the first time strong enough to approximate the *meanings of sentences*. To learn *what the world would have to be like* to conform to how a sentence (of a certain sort) represents it is to learn something approximating its meaning. The significance of this advance for the study of language can hardly be overstated. Having reached this stage, we had both a putative answer to the question *What is the meaning of a sentence?* and a systematic way of studying it.

This is roughly where the philosophically inspired study of linguistically encoded information stood in 1960. Since then, philosophers, philosophical logicians, and theoretical linguists have expanded the framework to cover large fragments of human languages. Their research program starts with the predicate calculi and is enriched piece by piece, as more natural-language constructions are added. Modal operators include *it is necessarily the case that*, *it could have been the case that*, and the counterfactual operator *if it had been*

*the case that* \_\_\_\_, *then it would have been the case that* \_\_\_\_\_. Operators involving time and tense can be treated along similar lines. Generalized quantifiers have been added, as have adverbs of quantification, and propositional attitude verbs such as *believe*, *expect*, and *know*. We also have accounts of adverbial modifiers, comparatives, intensional transitives, indexicals, and demonstratives. At each stage, a language fragment for which we already have a truth-theoretic semantics is expanded to include more features found in natural language. As the research program advances, the fragments of which we have a good truth-theoretic grasp become more powerful and more fully natural language-like. Although there are legitimate doubts about whether all aspects of natural language can be squeezed into one or another version of this representational paradigm, the prospects of extending the results so far achieved justify optimism about eventually arriving at a time when vastly enriched descendants of the original systems of Frege and Russell approach the expressive power of natural language, allowing us to understand the most basic productive principles by which information is linguistically encoded.

This, in a nutshell, is the dominant semantic conception in theoretical linguistics today. If all that remained were to fill in gaps and flesh out empirical details, philosophers would have done most of what was needed to transform their initial philosophical questions about mathematics into scientific questions about language. However, we haven't yet reached that point. While the dominant conception has made progress in using truth conditions to *model* representational contents of sentences, it has not paid enough attention to the demands that using and understanding language place on agents. Given the logical, mathematical, and philosophical origins of the enterprise, it could hardly have been otherwise. When what was at stake was, primarily, the investigation of the logical, analytic, or necessary consequences of mathematical and scientific statements,

there was no theoretically significant gap to be considered between what a sentence means and the claim it is used to make, and hence no need either to investigate how speaker-hearers might fill such gaps or to study what understanding and using a language consist in, and no need to individuate thoughts or meanings beyond necessary equivalence.

There have, to be sure, been important attempts to address these issues as the dominant semantic model has extended its reach beyond the formal languages of logic, mathematics, and science. We need, for example, to look no further than David Kaplan's *logic of demonstratives*, to find a way of accommodating the idea that what a sentence means and what it is standardly used to say are—though systematically related—not always the same. What we don't find in Kaplan, or in the dominant approach generally, is any retreat from the idea that advances in the understanding the semantics of natural language are closely and inextricably tied to advances in extending the reach of the methods of formal logic and model theory. This, I believe, must change if we are to reach our goal of founding a truly scientific study of language and information.

In this book, I will outline three steps in that direction. First, I will use examples involving several linguistic constructions to argue that we must stop oversimplifying the relationship between the information semantically encoded by (a use of) a sentence (in a context), on the one hand, and the assertions it is there used to make, the beliefs it is there used to express, and the information there conveyed by an utterance of it, on the other. It has often been assumed that the semantic content of a sentence is identical, or nearly so, with what one who accepts it thereby believes, and with what one who utters it thereby asserts. This is far too simple; there is a significant gap between the semantic contents of sentences and the information contents of their uses.

Second, I will argue that we need to pay more attention to what *understanding* a linguistic expression E requires—beyond, or other than, *knowing of* the representational content of E that it is the content of E. It is often assumed that since meaning is semantically encoded information, and since the information encoded by a nonindexical sentence S is the proposition p it expresses, understanding S is knowing of S that it encodes p. I will argue that this is not so. Semantic knowledge of this simple representational sort is insufficient for understanding because, as I will illustrate in chapter 4, to understand a word, phrase, or sentence is to be able to use it in expected ways in communicative interactions with members of one's linguistic community, which involves graded recognitional and inferential ability that often goes well beyond a cognitive grasp of content.<sup>3</sup> The semantic knowledge in question is also unnecessary for understanding a sentence because, as I will argue in chapters 2 and 4, to understand S is to be disposed to use S to entertain p—which, contrary to what is often assumed, doesn't require being disposed to make p the object of one's thought, or to predicate any relation holding between S and p. Once we have a proper understanding of what propositions really are, it will be easy to see that to entertain one is *not* to have any thought about or cognition of it at all, but to perform the cognitive operations in terms of which the proposition is defined.

This brings me to the final, and most foundational, change in the theoretical framework needed in our quest for a truly scientific study of language and information. Up to now, theorists have identified the semantic content of a sentence with information that represents the world as being a certain way, but they haven't yet given a plausible story about what such a piece of information is, whether

<sup>3</sup> This point is developed at length by my student Brian Bowman in his USC dissertation (2012).



linguistically encoded or not. This is our most urgent task, and the one on which I will concentrate most.

The currently dominant semantic approach correctly maintains that *S* represents things as being a certain way, and so has truth conditions, *because* the information *S* encodes—the proposition it expresses—represents things that way, and so has truth conditions. However, it misidentifies those propositions as functions from possible world-states to truth values, or, more simply, as sets of possible world-states (or other truth-supporting circumstances). However, neither such functions nor such sets can play the four roles typically demanded of propositions: (i) the primary bearers of truth and falsity, (ii) the objects of belief, assertion, and other attitudes, (iii) the contents of perceptual and cognitive states, and (iv) the meanings of (some) sentences. There are three main reasons that support this negative conclusion.

First, both functions from world-states to truth values and sets of such world-states are too coarse-grained to be meanings of sentences or objects of the attitudes. Worse, the strategies cooked up to mitigate the problem—(a) substituting finer-grained truth supporting circumstances for world-states, (b) developing a discourse model that substitutes so-called “diagonal propositions” for propositions semantically expressed as objects of crucial assertions, and (c) using two-dimensional semantic theories to assign pairs of coarse-grained propositions to sentences—have all failed to solve the problems they were supposed to address.<sup>4</sup>

Second, coarse-grainedness aside, neither sets of world-states nor functions from such to truth values can be meanings of sentences, while also being primary bearers of truth conditions. Meanings are

<sup>4</sup> Regarding (a), see Soames (1987, 2008); regarding (b), see Soames (2006c); regarding (c) see Soames (2005c).

the interpretations of sentences rather than entities that themselves require interpretation. But without interpretation by us, a set of world-states doesn't have truth conditions at all. In and of itself, such a set doesn't represent anything as being one way rather than another, and so doesn't impose any conditions the world must satisfy if the putative proposition is to be true. Consider the set containing just world-states 1, 2, and 3. Is it true or false? Since the set doesn't represent anything as being one way or another, it can't be either. We could, if we wanted, *interpret* it as representing *the actual world-state as being in the set*, and so as being true iff no world-state *outside* the set were the unique world-state that is instantiated. But we could equally well interpret it as representing the actual world-state as *not being in the set*, and so as being true iff no world-state *inside* it is instantiated. Without interpretation by us, the set doesn't represent anything, and doesn't have truth conditions. Since propositions aren't things we interpret, but are themselves the interpretations we give to sentences, they aren't sets of world-states.

The *function* that assigns *truth* to some world-states and *falsity* to others is no better. Suppose we replaced truth and falsity with 1 and 0. What does a function that assigns 1 to some states and 0 to others represent? Without interpretation by us, it doesn't represent anything. Nor, as indicated by (i)–(iv), does it help to appeal to truth and falsity rather than to 1 and 0.

- (i) Truth is the property that a proposition has when the world is as the proposition represents it to be, as well as a property which, when predicated of a proposition *p*, gives us a claim—that *p is true*—that one is warranted in accepting, asserting, rejecting, denying, believing, or doubting iff one is warranted in taking the same attitude to *p*. Since this is what truth is, propositions are conceptually prior to truth,

in which case truth can't be one of the things from which propositions are constructed.

- (ii) Even if this weren't so, the function that assigns world-states 1–3 truth, and all others falsity, doesn't, in and of itself, represent anything as being one way rather than another—any more than does the set of world-states of which it is the characteristic function. The illusion that it does comes from telling ourselves a story that relies on a different, conceptually prior, notion of propositions. In this story, which depends on thinking of world-states as properties that can be predicated of the universe, each assignment of a truth value to a world-state  $w$  is correlated with *the proposition that predicates  $w$  of the universe*—which is thereby true (false) iff the universe is (isn't) in state  $w$ . A function from world-states to truth values can then be associated with the (possibly infinite) disjunction of the propositions correlated with its assignments of truth to world-states. However, far from vindicating the idea that such functions *are* propositions, it presupposes an antecedent conception of propositions according to which they are *not* such functions.
- (iii) The conception of propositions as functions from world-states to truth values goes hand in hand with a conception of properties as functions from world-states to extensions. This flies in the face of taking the “worlds” of possible-worlds semantics to be properties—for surely a world-state *isn't* a function from world-states to anything. But if properties have to be taken as basic, rather than explained away, surely propositions should have the same status.
- (iv) It is better *not* to take world-states as primitive either. The best account takes them to be properties of making complete world-stories (the constituents of which are propositions) true. Thus, *both truth and world-states* are conceptually

downstream from propositions, and so are not the conceptual building blocks from which propositions are constructed.<sup>5</sup>

Finally, the proponent of possible-worlds semantics faces a dilemma. Suppose, as is common, the theorist takes the two-place predicate  $x_s$  is true at  $y_w$  (where the variable ‘ $x_s$ ’ is assigned a sentence  $S$  and the variable ‘ $y_w$ ’ is assigned a world-state  $W$ ) to be an undefined primitive. Then, the theorist has no way of answering the question, “What, if anything, does the theorem *For all world-states  $y_w$ , ‘Mi libro es rojo’ is true at  $y_w$  iff at  $y_w$ , my book is red*, tell us about the meaning of the Spanish sentence?” By contrast, suppose one understands  $x_s$  is true at  $y_w$  (relative to an assignment of  $S$  and  $W$  to the respective variables) as telling us that if  $W$  were instantiated then *the proposition that  $S$  actually expresses* would be true. To understand ‘is true at’ in this way is to presuppose antecedent notions of *the proposition  $S$  expresses* and *the monadic notion of truth* applying to it. Taking these antecedent notions at face value, the theorist can use the commonsense truism—[if ‘ $S$ ’ means, or is used to express, the proposition that  $P$ , then necessarily the proposition expressed by ‘ $S$ ’ is true iff  $P$ ]—together with the truth-conditional theorem—*For all world-states  $y_w$ , ‘Mi libro es rojo’ is true at  $y_w$  iff at  $y_w$ , my book is red*—to conclude that the sentence ‘Mi libro es rojo’ means (expresses) something necessarily equivalent to the proposition that my book is red.

Although taking this route doesn’t identify what the sentence means, it does allow us to extract substantial information about its meaning from a statement of its truth conditions at possible world-states. The price of this happy result is the obligation to explicate the *prior notions of the proposition expressed by  $S$  and the monadic notion of truth* applying to it. Since possible-worlds semanticists have typically refused to acknowledge this price, let alone to pay it, they are in no

<sup>5</sup> See Soames (2007a) and chapter 5 of (2010a).

For general queries, contact [webmaster@press.princeton.edu](mailto:webmaster@press.princeton.edu)

position to claim that their theories provide any information at all about meaning. By contrast, those of us who wish to preserve the great progress in the study of language made by applying the methods of intensional semantics to natural language must find a way of paying the price for them by explaining what propositions really are and how they can be added to possible-worlds semantic theories as genuine truth-condition determiners. This is our most serious problem.

It is tempting to think that it can be solved by returning to traditional Fregean or Russellian conceptions of structured propositions. However, the utility of such conceptions is severely limited. Taking their cues from these traditional theories, many contemporary proponents of structured propositions address the coarse-grainedness problem by using n-tuples of objects and properties to model propositions. Although this approach does a better job of individuating the objects of the attitudes than does any conception of propositions as constructions of truth-supporting circumstances, contemporary conceptions of structured propositions don't go far enough. As I will argue at length in later chapters, they, like possible-worlds conceptions of propositions, wrongly foreclose needed analyses of the assertions made, and beliefs expressed, by many utterances, including, most obviously, utterances of propositional attitude ascriptions. Worse, the n-tuples standardly employed by contemporary proponents of structured propositions are merely models of the real things. Since n-tuples, or any other purely formal structures of objects and properties, don't, without interpretation by us, represent the world as being one way rather than another, they can't be meanings or primary bearers of truth conditions. The same is true of the propositions originally proposed by Frege and Russell, as attested by recent work on the so-called *problem of the unity of the proposition*, which undermined them.<sup>6</sup>

<sup>6</sup> See Soames (2010c); Soames (2014b), chapters 2, 3, 7, and 9; and King, Soames, and Speaks (2014), chapter 3.

The concern underlying the so-called unity problem was also the basis for Donald Davidson's most telling objection to semantic theories that postulated structured propositions as meanings of sentences. Commenting on traditional conceptions of propositions in 1967, he aptly remarked:

Paradoxically, the one thing meanings do not seem to do is oil the wheels of a theory of meaning—at least as long as we require of such a theory that it non-trivially give the meaning of every sentence in the language. My objection to meanings in the theory of meaning is not that they are abstract or that their identity conditions are obscure, but that they have no demonstrated use.<sup>7</sup>

His complaint was that taking structured propositions to be the meanings (or, in more recent terminology, *semantic contents*) of sentences (relative to contexts) doesn't help in constructing a theory of meaning, unless one can read off how a sentence represents things to be from the specification of the structured proposition it expresses. Since this crucial information *can't* be read off Fregean, Russellian, or any other traditional account of structured propositions, a new conception of propositions is needed.<sup>8</sup>

To provide such a conception, while retaining the insights of Frege and Russell, we must reverse their explanatory priorities. Instead of deriving the intentionality of agents from independently representational propositions, we must explain the intentionality of propositions in terms of the conceptually prior ability of agents to represent the world in thought and perception. This will be my task in the next chapter.

<sup>7</sup> Davidson (2001), quoted at pp. 21–22.

<sup>8</sup> See Soames (2010c), pp. 49–55, for discussion.