

1

Worldly Semantics and the Myth of the Given

1.1 Introduction

In this opening chapter, I lay out what I take the explanatory aim of a semantic theory to be, and I introduce “worldly semantics” as a strategy for accomplishing this aim. I’ll then state the thesis, which I’ll defend in the next two chapters, that worldly semantics is committed to an instance of what Wilfrid Sellars calls “the Myth of the Given,” spelling out at some length what I take this term to pick out. Finally, I’ll clarify the target by drawing a distinction between “explanatory” and “elucidatory” models in semantics and illustrating what is at stake in locating, for instance, possible worlds semantics on one side of that distinction rather than the other.

1.2 Our Semantic Aim

Language speakers aren’t mindless automata. Generally, they know what they’re saying when they use expressions of a language that they know how to speak. They have this knowledge because they know what these expressions mean. The aim of semantics is to understand what it is in which this aspect of the capacity to speak a language, knowledge of meaning, consists. Here is one particularly clear statement of this aim from Seth Yalcin (2018):

I take it that in natural language semantics, the aspect of reality we are seeking some understanding of is a dimension of human linguistic

competence—informally, knowledge of meaning. Competent speakers of a language know ('cognize', etc.) the meaning features of expressions of their language. The semanticist is interested in modeling this state of mind and the associated semantic features, (2018, 353).

I take it that most semanticists working in the Chomskian tradition of generative grammar think of the discipline roughly along these lines.¹ In semantics, we are aiming to understand the knowledge of meaning that competent speakers have. We take it that this knowledge explains certain aspects of their linguistic behavior—the behavior that they exhibit in virtue of knowing what expressions of their language mean—and what we're aiming to do, in constructing a semantic theory, is to explain this behavior by modeling the knowledge of meaning that accounts for it. The main task of the semantic theorist is to assign *semantic values* to the expressions of the language for which she is constructing a semantic theory. These semantic values are the entities in the semantic theory that are meant to be mathematically defined models of the meanings of the expressions of which speakers of that language have knowledge. So, meanings are theorized to play a certain sort of explanatory role: knowledge of them is taken to explain certain aspects of speakers' behavior. And semantic values are entities in the theorist's model that are mathematically defined in such a way that they satisfy certain properties, properties which are either identical or structurally analogous to the properties that the theorist takes it that meanings must have in order for them to play the explanatory role that they are theorized to play.

Now, there are several properties that the meanings of sentences are taken to have that are meant to be modeled by the assignment of semantic values to them. To limit the scope of my discussion, I will focus here on just one crucial such property. The meanings of sentences are taken to determine facts consisting

¹For instance Gennaro Chierchia and Sally McConnell-Ginnet (1990) write "It is the application of mathematical models to the study of the cognitive phenomenon of linguistic knowledge that most generative linguists recognize as their aim," (2).

in these sentences standing in relations of entailment and (in)compatibility (or, consequence and (in)consistency) to one another. One important role of semantic values is to model meanings in such a way that we can explain these facts, thereby explaining speakers' knowledge of them, and thereby explaining their behavior that is a manifestation of this knowledge. Here is Yalcin again, in a different paper of his, stating this point:

[S]emantic values are assumed to be the sorts of things consequence and consistency relations are articulated in terms of: when $\Gamma \vdash \varphi$ holds, this is (at least partly) because of the semantic values of (the sentences in) Γ and of φ , respectively. Hypotheses about semantic values can thereby serve to predict, and ground, entailment and consistency facts, hence knowledge of such facts, (2014, 24).

Assigning semantic values to sentences, and then articulating consequence and consistency relations in terms of these semantic values, we model the meanings of sentences in such a way that we are able to explain, with the use of our model, how facts consisting in sentences entailing or being incompatible with one another obtain in virtue of these sentences meaning what they do. Then, by modeling of speakers' knowledge of the meanings of these sentences as knowledge of the semantic values we assign to them, our theory will explain how their knowledge of entailment and incompatibility relations obtaining between sentences, knowledge which explains certain aspects of their behavior, is determined by their knowledge of the meanings of these sentences.

1.3 A Toy Language

To see more determinately what aspects of speakers' behavior we're trying to explain in assigning semantic values to sentences in terms of which entailment and incompatibility relations can be articulated, it will be helpful to introduce a very simple "toy language" and then consider what theoretical work a semantic

theory for this toy language should be able to do insofar as it aspires to this explanatory aim. So, imagine a small linguistic community whose members speak a language consisting of the following expressions:

1. Three names: “*a*,” “*b*,” and “*c*”
2. Three 1-place predicates: “is white,” “is gray,” and “is black”
3. Three 2-place predicates: “is lighter than,” “is darker than,” and “is the same shade as”
4. One unary sentential operator: “It is not the case that”
5. Two binary sentential operators: “and” and “or”
6. Left and right parentheses (to avoid ambiguity): “(” and “)”

This is their basic vocabulary: the set of simple expressions that they are able to employ. The grammar of their language, through which complex expressions can be constructed from these simple ones, can be recursively specified as follows:²

1. Any name followed by a 1-place predicate is a sentence.
2. Any name followed by a 2-place predicate and then another name is a sentence.
3. If φ is a sentence and U is a unary operator ($U\varphi$) is a sentence.
4. If φ and ψ are sentences and B is a binary operator ($\varphi B\psi$) is a sentence.
5. If some string of lexical items can't be constructed by the use of these rules, it's not a sentence.

Call any sentence that contains no sentential operators an “atomic sentence.” There are thirty-six atomic sentences of our toy language, including, for instance “*a* is white,” “*b* is darker than *c*,” “*c* is gray,” and so on. There are infinite non-atomic sentences, formed by conjoining atomic sentences with operators and parentheses. Our toy language consists in this infinite set of sentences. This is, of course, a woefully impoverished language, and it can hardly be called a language at all, but it is enough of a language for our purposes here.

²I am using “grammar” here not in the Chomskyan sense.

I will make extensive use of this toy language throughout this dissertation, so it is worth saying a few words now to preliminarily justify my doing so. As it is probably clear, I have laid out the simplest toy language that I possibly could. I hope it will be clear in what follows that I could have introduced a much fancier toy language here—with more vocabulary belonging to the grammatical types introduced here, vocabulary belonging to additional grammatical types, and a more complex grammar to accommodate this additional vocabulary—and used it to the same end in the next few chapters. I take it that doing this would have only unnecessarily complicated things, so that is why I have not done so, saving the introduction of more sophisticated toy languages to the positive part of the dissertation where I develop the alternative framework of discursive role semantics.³ Of course, it takes some cognitive dissonance to imagine that we could really have what Brandom calls an “autonomous discursive practice” whose members speak this “language” and it alone. Indeed, as we’ll later see, there could not be such a practice, and we will need a richer language in order to be able to think of there as being speakers who employ that language and it alone. For now, however, let us engage in the imaginative exercise of taking there to be speakers who speak this language and it alone, grasping the meanings of the expressions that belong to it and behaving certain ways in virtue of grasping these meanings.

Suppose our speakers act in such a way that shows that they take the sentences “*a* is darker than *b*” and “*b* is lighter than *a*” to be synonymous. Now, if they have semantic vocabulary, they might say “These two sentences are synonymous,” “These two sentences mean the same thing,” or “One says the same thing in

³It is also worth pointing out that, grammatically, I have taken some shortcuts and treated this toy language in such a way that much more closely resembles the formal language of first-order logic than a natural language like English. Once again, this is just for simplicity, and nothing important hangs on this. Contemporary work in semantics, following Montague (1974), “reject[s] the contention that an important theoretical difference exists between formal and natural languages,” (188).

uttering either of these two sentences," but we need not even credit them with this sort of vocabulary in order to get our basic explanandum into view; it is sufficient that they, in their linguistic practices, treat the two sentences in the way that two synonymous sentences ought to be treated. For instance, whenever a competent speaker utters one, they'll be prepared to utter the other, if an incompetent speaker utters one but refuses to utter the other, they'll be corrected by competent speakers, and so on. Similarly, competent speakers of this language take the sentences "*a* is black" and "*b* is gray" to jointly entail the sentence "*a* is darker than *b*." Any competent speaker that utters both "*a* is black" and "*b* is gray" will also be prepared to utter "*a* is darker than *b*," and if an incompetent speaker utters the first two but refuses to utter the third, they'll be corrected, and so on. Finally, they take the sentences "*a* is black" and "*a* is gray" to be incompatible. They'll never utter both sentences at the same time, they'll correct incompetent speakers that do, and so on. These activities, we theorize, are manifestations of their knowledge of the meanings of the sentences "*a* is darker than *b*," "*b* is lighter than *a*," "*a* is black," "*b* is gray," and "*a* is gray." That is to say, it is in virtue of knowing what these sentences mean that the speakers of our toy language behave in these ways. What we want to do, in constructing a semantic theory for their language, is understand this knowledge of meaning in such a way that enables us to explain this behavior. Officially, what we want to do is assign semantic values to these sentences, formal models of their meanings, such that if speakers know that these sentences have these semantic values, they'll know these sentences stand in these semantic relations, since, if they know that these sentences stand in these semantic relations, they'll behave in these ways.

Given this explanatory aim, assigning semantic values to sentences should enable us to account for facts like the following:

- F1. The sentence "*a* is darker than *b*" is synonymous with the sentence "*b* is lighter than *a*."

- F2. The sentences “*a* is black” and “*b* is gray” jointly entail the sentence “*a* is darker than *b*.”
- F3. The sentence “*a* is black” is incompatible with the sentence “*a* is gray.”

If, by assigning semantic values to the sentences “*a* is darker than *b*,” “*b* is lighter than *a*,” “*a* is black,” and “*b* is white,” we are able to account for these facts, then, by thinking of speakers’ knowledge of the meaning of these sentences in terms of their knowledge of these semantic values, we can explain their knowledge of these facts, and, accordingly, the behavior they exhibit in virtue of having this knowledge.

1.4 The Meanings of Content Words

In specifying (F1)-(F3), I have picked out by way of example only one class of synonymy, entailment, and incompatibility relations that obtain between sentences of this toy language: the class of *material* rather than *formal* relations of synonymy, entailment, and incompatibility. For instance, the sentences “*a* is gray” and “*a* is black” are *materially* incompatible, whereas the sentences “*a* is gray” and “It’s not the case that *a* is gray” are *formally* incompatible. Articulating this distinction with the use of a more contemporary vocabulary, material semantic relations are relations that obtain between sentences in virtue of the meanings of (what are often called) the “content words” contained in those sentences, words like “gray,” “black,” or “darker than,” whereas formal semantic relations are relations that obtain between sentences in virtue of the meanings of (what are often called) the “function words” like “not,” “and,” and “or.”⁴ Now, articulating exactly what

⁴See, for instance, Lobner (2002, 4-5) and Gendler Szabo (2019) for an articulation of this distinction with the use of this terminology. This terminology is somewhat confusing, since the meanings of content words are themselves taken to be functions. Kearns (2011) uses the terms “categorematic” and “syncategorematic,” but this is also potentially problematic, given that the meanings of those terms, in a contemporary context, don’t directly map on to their classical usage.

this distinction consists in will depend on the sort of semantic theory that one ends up endorsing. In almost every semantic theory, however, the semantic values assigned to content words will form the foundation on the basis of which the rest of theory will be constructed.⁵

It is only given the assignment of semantic values to these simple content words that the assignment of semantic values to function words, generally conceived in terms of operations on the semantic values of content words, makes any sense at all. For instance, in a possible worlds semantics, semantic values for logically complex sentences can be understood in terms of the set-theoretic operations of complementation, intersection, and union only insofar as atomic sentences are assigned sets of possible worlds as semantic values, and atomic sentences can be assigned sets of possible worlds as semantic values only insofar as the content words that make them up, words like “gray” or “black,” are assigned suitable semantic values, for instance, functions that map each possible world to the set of things that are gray in that world or black in that world. So the assignment of suitable semantic values to content words is required at the base level of semantic theories. This fact about the structure of semantic theories follows directly from such a theory’s commitment to the compositionality of meaning: that the meaning of a complex sentence is determined by the meaning of its parts and the way those parts are put together. If we cannot think of the meanings of content words as adequately modeled by the semantic values that a compositional semantic theory assigns to them, the whole theory that is based on these basic assignments falls like a house of cards.

Despite the fact that semantic theories require the assignment of semantic values to content words at their base level, most semanticists do not concern themselves with these basic assignments of semantic values. While it does fall

⁵This is, I take it, implicitly acknowledged by almost all semantic theorists; one theorist who is explicit about this is Gendler Szabo (2019).

to the semantic theorist to specify the semantic types corresponding to content words of different syntactic categories, the task of specifying the meanings of these basic expressions in any substantive way is not a task for semantics, properly construed. Distinctions in meaning between such words as “gray” and “black,” insofar as they belong to the same syntactic category, are, “from the point of view of semantic theory, simply brute,” (Yalcin 2018, 350). So, for instance, a possible worlds semantics might assign to the 1-place predicate “gray” the function that maps each world to the set of things that are gray in that world, and it will assign “black” the function that maps each world to the set of things that are black in that world. Of course, such a theory won’t tell us what it is for something to be gray as opposed to white, but we shouldn’t expect it to. A dictionary can tell us this. To think that it’s the job of the semantic theory to tell us what a dictionary would tell us would be to confuse semantics with lexicography, and that, as Richard Thomason (1975) says, is “a persistent and harmful source of misunderstanding in matters of semantic methodology,” (48). Yalcin quotes this sentiment, expressed by Thomason in his introduction to Montague’s *Formal Philosophy*, in support of this attitude towards the meanings of content words:

[W]e should not expect a semantic theory to furnish an account of how any two expressions belonging to the same syntactic category differ in meaning . . . ‘Walk’ and ‘run’, for instance, and ‘unicorn’ and ‘zebra’ certainly do differ in meaning, and we require a dictionary of English to tell us how. But the making of a dictionary demands considerable knowledge of the world [of a sort the semantic theorist should not be expected to furnish],” (Yalcin 2018, 350, quoting Thomason (1975); Thomason’s italics, Yalcin’s bracketed addition.)

Knowing the meanings of content words like “walk” and “run” or “black” and “gray.” requires “considerable knowledge of the world.” A semantic theorist, in taking speakers to have knowledge of the meanings of words like “walk” and “run” or “black” and “gray,” appeals to this worldly knowledge that speakers have—their knowledge of what it is for something to walk as opposed to run,

or what it is for something to be black as opposed to gray—but this worldly knowledge, which is an ingredient in speakers' knowledge of meaning, is to be distinguished from the properly semantic knowledge that is the proper concern of the semantic theorist. As such, it is sufficient for the semanticist to say something along the following lines:

The meaning of the predicate "gray" is determined entirely (or, at least, sufficiently for our purposes) by the fact that it is correctly applied to some object just in case that object is gray. Accordingly, we can model the meaning of the predicate "gray" as a function that maps each possible world, each way for things to be, to the set of things that are gray in that world; the set of things to which that predicate is correctly applied. What we model in modeling the meaning of this predicate in this way is what a speaker knows in knowing the meaning of this predicate; they know that, however things are, this expression is to be applied to something just in case that thing is gray.

Having given this justification of their formal model of meanings of 1-place predicates such that they compose in the right ways with the meanings of other types of expressions, the semanticist can leave it to the lexicographer to say, substantively, what it is for something to be gray, how being gray differs from being black, and so on.

This apparent division of labor may seem to be of a piece with the divide-and-conquer methodology found throughout the natural sciences. However, implicit in this way of thinking about speakers' knowledge of meaning is the idea that speakers' knowledge of certain worldly facts—for instance, the fact that the property of being gray is incompatible with the property of being black (i.e. being gray is a way for something to be such that, if something is that way, it cannot be black)—is explanatorily prior to their knowledge of certain semantic ones—for instance, the fact that the predicate "gray" is incompatible with the predicate "black." It is because the former sort of knowledge, the worldly knowledge, is not taken to be the proper object of a semantic theory that the knowledge of the

incompatibility of the predicates, from the point of view of the semantic theory, can be taken to come for free as a direct consequence of semantic values for content words that are “from the point of view of semantic theory, simply brute,” (Yalcin 2018, 350). For instance, in a possible worlds semantics, one simply assigns the predicates intensions that are assumed to be disjoint, appealing to one’s own knowledge of what it is for something to be gray or black in the assessment of these intensions as disjoint. This appeal to one’s own knowledge of what it is for something to be gray or black can be taken to be unproblematic only insofar as this knowledge that one appeals to is taken to be “knowledge of the world of a sort the semantic theorist should not be expected to furnish,” (Yalcin 2018, 350). This worldly knowledge is taken to underlie the semantic knowledge that constitutes of the base of the semantic theory, the knowledge of the meanings of content words like “gray” and “black” that grounds knowledge of facts such as (F1)-(F3). The semantic theories I will concern myself with in the negative part of this dissertation all assign semantic values to content words in accord with this theoretical orientation. They are all versions of what I will call “worldly semantics.” Let me now lay out, in abstract terms, the basic structure of a worldly semantic theory.

1.5 The Basic Structure of Worldly Semantics

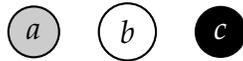
Most work in contemporary semantics is guided by the following core idea, which I’ll quote directly from the introductory textbook in formal semantics by Dowty, Wall, and Peters (1981):

To know the meaning of a (declarative) sentence is to know what the world would have to be like for the sentence to be true, (4).⁶

⁶See also the widely used introductory textbook by Heim and Kratzer (1998), where they open with the sentence, “To know the meaning of is to know its truth conditions,” going on to tell us, to know the meaning of a sentence, you don’t have to know whether it is true; “What you do know, however, is what the world would have to be like for it to be true,” (1).

To see how this core idea applies to the speakers of our toy language, let's suppose that there are only three things in the world in which they live—*a*, *b*, and *c*—and that these are the things that are named by the names “*a*,” “*b*,” and “*c*.” Furthermore, let's suppose that there are only three ways that these three things can be—completely white, completely gray, or completely black—and that these are the ways the speakers of our toy language say that something is when they say of it that it “is white,” “is gray,” or “is black.” Finally, let's suppose that there is only one shade of gray, so if something is darker or lighter than something else, it can't be the case that they're both the same shade.

Now, consider the sentence “*a* is gray.” There are many ways that the world of the speakers of our toy language can be such that this sentence is true. The world can be such that *a* is gray, *b* is white, and *c* is black. That is, the world can be like this:



Alternately, the world can be such that *a* is gray, *b* is gray, and *c* is white. That is, the world can be like this:



As far as the truth of “*a* is gray” is concerned, whether *b* is white or *b* is gray or whether *c* is black or *c* is white does not matter. What does matter is whether or not *a* is gray. The sentence “*a* is gray” is true just in case the world is such that *a* is gray. To know the meaning of this sentence, on a truth-conditional theory of the sort described by Dowty, Wall, and Peters, is to know just this.

Now, there are different ways in which this basic idea, expressed by Dowty, Wall, and Peters, can be implemented in a semantic theory in order to arrive at formally specifiable semantic values that are meant to serve as models of

what speakers know in knowing the meaning of a sentence. I will consider what I take to be the two basic ways in which this idea can be implemented in the next two chapters. For now, however, let's consider the general structure of a theory that conforms to this basic idea. Such a theory will be a version of what I will call "worldly" semantics. On a worldly semantic theory, we take speakers' knowledge of meaning to asymmetrically depend on their knowledge of "worldly" entities and their relations. The general sorts of "worldly" entities to which a worldly semantics appeals might be picked out with expressions such as "objects," "properties," "relations," "states of affairs," "possible worlds," and so on. Which of these sorts of entities are given priority over the others will vary from theory to theory, but, in a worldly semantics, speakers are taken to have knowledge of entities of these sorts and knowledge of relations that entities of these sorts stand to one another, and their knowledge of the meanings of sentences of their language is taken to depend on this worldly knowledge. To see how a worldly semantics is supposed to work, consider the sort of explanation of (F3) specified above—the fact that the sentence "*a* is gray" is incompatible with the sentence "*a* is black"—that we would provide if we endorse a worldly semantics.

We start by taking the speakers of our toy language to have a bit of worldly knowledge: They know that if *a* is gray, then it can't be the case that *a* is black. This bit of worldly knowledge might be analyzed in different ways. We might analyze it in terms of ways the world as a whole can be, saying that speakers know that, however the world can possibly be, the set of things that are gray in the world and the set of things that are black in the world are disjoint—they don't have any elements in common. Spelling this out a bit, we might take our speakers to have a grip on a space of different possible ways for the world as a whole to be, a space of different "possible worlds," and know that, for each point in this space, each possible world, the set of things that are gray and the set of things that are black are disjoint—these two sets do not have any common elements. So, whichever

element of the total set of possible worlds is actual, if a is an element of the set of gray things, then it isn't an element of the set of black things. This is one way to analyze what it is that our speakers know in knowing that, if a is gray, then it can't be the case that a is black. Since the worldly entities of which speakers are taken to have knowledge are worlds as a whole, I'll call it an "extra-worldly" analysis. Alternately, we might analyze our speakers' knowledge of the fact that if a is gray, then it can't be the case that a is black entirely in terms of ways things in this world can be. We might say that our speakers know that no single thing can, at the same time, be both gray and black. Spelling this out a bit, we might say that there is a basic modal fact about the property of being gray and the property of being black, one that obtains in virtue of the essences of these two properties; it is not possible for a single thing to, at one time, instantiate both the property being gray and the property of being black. So, since a is a single thing, if a instantiates the property of being gray, it is not possible for it, at the same time, to instantiate the property of being black. This is a different way to analyze what it is that our speakers know in knowing that, if a is gray, then it can't be the case that a is black. Since the basic worldly entities of which speakers are taken to have knowledge here are things *in* the world, objects and properties that these objects might have, I'll call it an "intra-worldly" analysis. However we want to analyze our speakers' worldly knowledge of the fact that if a is gray, it can't be the case that a is black, if we endorse a worldly semantics, we'll think of our speakers' knowledge of the fact that the sentence " a is gray" is incompatible with " a is black" as asymmetrically depending on a bit of worldly knowledge of this sort.

Consider first the semantic picture suggested by an extra-worldly conception of this bit of worldly knowledge. On a standard variant of extra-worldly semantics, the semantic value of an expression is a function that maps possible worlds to extensions. For names, these extensions are taken to be particular objects, and for 1-place predicates, these extensions are taken to be sets of objects. So, the semantic

value of “ a ” is a function that maps each possible world to a , a particular thing that we may assume exists in each world, and the semantic value of “is gray” is a function that maps each possible world to the set of things that are gray in that world.⁷ Now, we have a rule of composition that says that, for any sentence of the form “ n is F ,” consisting in a name “ n ” concatenated with a 1-place predicate “is F ,” the semantic value of that sentence is the set of worlds w such that the object to which the semantic value of “ n ” maps w , namely n , is an element of the set of objects to which the semantic value of “ F ” maps w , the set of things that are F in w . So, the semantic value of “ a is gray” is the set of worlds in which a is an element of the set of gray things in that world. Likewise, the semantic value of “ a is black” is the set of worlds in which a is an element of the set of black things in that world. Now, if one knows that, for each possible world, there is no object that is an element of both the set of gray things in that world and the set of black things in that world, and one knows that “ a is gray” and “ a is black” have the semantic values that they do, one will know that the sets of worlds that are the semantic values of “ a is gray” and “ a is black” are disjoint. That, according to an extra-worldly semantics, is just what it is to know that the sentences “ a is gray” and “ a is black” are incompatible. In this way, an extra-worldly semantics takes speakers’ semantic knowledge to asymmetrically depend on a bit of extra-worldly knowledge.

Now consider the semantic picture suggested by an *intra*-worldly conception of this bit of worldly knowledge. On a standard variant of an *intra*-worldly semantics, the semantic value of a name is the object named by that name, and the semantic value of a 1-place predicate is the property expressed by that predicate. So, the semantic value of “ a ” is a , the object that is named by “ a ,” and the semantic

⁷This assumes, following Kripke, that names are “rigid designators.” At some point, we’d have reason to drop the assumption that a exists in each world. In which case, we can take the semantic value of a name to be a *partial* function that maps each possible world in which the object actually named by that name exists to that object. I’ll continue to make such simplifying assumptions here.

value of “gray” is the property of being gray, the property that is expressed by “gray.”⁸ Now, we have a rule of composition that says that, for any sentence of the form “*n* is *F*,” consisting in a name “*n*” concatenated with a 1-place predicate “is *F*,” the semantic value of that sentence is a structured proposition that represents the object that is the semantic value of “*n*” as instantiating the property that is the semantic value of “*F*.” So the semantic value of “*a* is gray” is a structured proposition that represents *a* as instantiating the property of being gray. Likewise, the semantic value of “*a* is black” is a structured proposition that represents *a* as instantiating the property of being black. Now, if one knows that the property of being gray and the property of being black are incompatible in the sense that it is not possible for a single object to instantiate both of these properties, one will know that these two propositions that are the semantic values of “*a* is gray” and “*a* is black” cannot both be true; taken together, they represent a single thing as being two ways that a single thing cannot, at a single time, be. That, according to an intra-worldly semantics of this sort, is just what it is to know that the sentences “*a* is gray” and “*a* is black” are incompatible. In this way, an intra-worldly semantics takes speakers’ semantic knowledge to asymmetrically depend on a bit of intra-worldly knowledge.

Though this theoretical structure is rarely as explicit as I have made it out here, I take it that worldly semantics is pretty much ubiquitous in contemporary theorizing about speakers’ knowledge of meaning in both philosophy and linguistics. Almost every working semanticist practices a variant of worldly semantics. Semantic theories that are often taken to be on opposite sides of fundamental dividing lines in semantic theorizing—for instance, truth-conditional semantics vs. dynamic semantics, possible worlds semantics vs. situation semantics—will

⁸This will need some refinement, depending on the variant of intra-worldly semantics we consider. Soames (2014, 2015), for instance, takes the semantic values of names and predicates not to be objects and properties themselves, but acts of cognizing objects and properties. Once again, I’ll make simplifying assumptions in the consideration of intra-worldly semantics, but nothing will hang on this.

generally still all be variants of worldly semantics.⁹ Nevertheless, I will argue here that such semantic theories are not able to do the explanatory work that a semantic theory is supposed to be able to do. They are not able to give us anything resembling an account of the aspect of semantic competence consisting in knowledge of meaning. The basic problem is this: we cannot give an account of speakers' knowledge of meaning as asymmetrically depending on their knowledge of worldly entities and their relations in the way that a worldly semantics requires us to do because this worldly knowledge itself depends on speakers' knowledge of meaning. To give a name to the general form of the problem of which this problem is a specific instance and to give some philosophical context for the line of critique I am about to prosecute, my claim is that, in proposing to explain speakers' knowledge of meaning in the way that they do, proponents of worldly semantics are guilty of a version of what Wilfrid Sellars calls "The Myth of the Given," (1956).

Now, my main aim in this dissertation is not the exegesis of Sellars.¹⁰ I do take it, however, that all of my claims, both negative and positive, are basically Sellarsian ones. Accordingly, it is worth doing a little work to articulate, in Sellars's own terms, what the general form of the problem is and how, by his own lights, worldly semantics is an incarnation of it.

1.6 The Myth of the Given

Sellars's term "The Myth of the Given" has become something of a buzzword among contemporary philosophers who have taken themselves to have learned

⁹It's worth noting that dynamic semantics, of the sort proposed by Veltman (1996), do not conform to the basic truth-conditional dictum quoted above from Dowty, Wall, and Peters. However, they nevertheless require conceiving of extra-worldly knowledge as underlying knowledge of meaning. I'll address such theories explicitly in Chapter 4.

¹⁰I have a companion paper (Simonelli M.S.), of which a version of the next section is modified excerpt, where I do develop an exegesis of Sellars according to which it is clear that worldly semantics falls within the scope of the Myth of the Given.

some lesson from Sellars, and it has become a cause of frustration for some contemporary philosophers who aren't part of the club that throws this term around but get it thrown at them. David Chalmers (2010), who gets this term thrown at him as much as anyone in contemporary philosophy, takes the term to pick out the view that "experiences have a special epistemic status that renders them 'given' to a subject," going on to say, "Sellars's (deliberately abusive) term for the view has caught on, and today it is not uncommon for this label to be used in criticizing such views as if no further argument is necessary," (299). I am quite sympathetic to Chalmers's frustration here. It is quite common for the label "the Myth of the Given" to be applied to some view in order to dismiss it without any further argument. Indeed, in many cases of its application, it seems to function as a *mere* label, without any clear descriptive content. The first thing I want to point out is that the set of views that are aptly characterized as instances of "the Myth of the Given" extends much more widely than specifically views about sensory experience. Sellars does start his critique of the Myth of the Given by considering views in which sensory experiences have a special epistemic status, but he takes this to be only "a first step in a general critique of the entire framework of givenness," (1956/1997, 14). When I use the term, I mean to speak about this general framework.

There is considerable debate among commentators as to what, exactly, the general framework of givenness is. Some commentators, such as deVries and Triplett (2000), take the Myth to involve a rather particular version of foundationalism. On this way of construing things, Sellars's critique of the Myth would rather parochially pertain to views in epistemology that were prevalent in Sellars's day but are widely disregarded nowadays. As such, it'd be hard to see, on such a construal that the views in semantics that I am attacking here, which seem rather far removed from the epistemological foundationalism of the early part of the 20th century, could be aptly characterized as instances of the Myth. Other

commentators, such as Brandom (1997), construe the Myth as the view that there could be a form of non-conceptual awareness that directly entails having conceptual knowledge. Once again, on this way of construing the Myth, it pertains to a rather particular view in the philosophy of mind (and, indeed, one that few philosophers of mind nowadays would accept), and it's hard to see how it could pertain to the views in semantics that I am attacking here. So, if the Myth of the Given is not to be identified as deVries and Triplett or Brandom identify it, how is it to be identified? What *is* the Myth of the Given?

The answer to this question, I think, is surprisingly straightforward; the term "Myth of the Given" actually functions as perfectly sufficient as a description of what it picks out. The Myth of the Given is simply any conception of our knowledge of some aspect of reality as simply *given* to us, and intelligible only as given in this way. The basic problem with such a conception—what makes any such conception a *myth*—is that, by thinking of knowledge of some aspect of reality as given in this way, we preclude ourselves from thinking of our knowledge as rational, and thus, as genuinely knowledge. Holding something rationally requires being able, at least in principle, to put it in to question and, in response to that question, articulate the reasons for holding it. If knowledge of some aspect of reality is taken to be simply given, and intelligible only as such, then this knowledge constitutes a stopping point in the inquiry into our knowledge of reality, at which no questions can be asked. But if no questions can be asked, then no reasons can be given, and so we cannot make sense of our knowledge of the aspect of reality that is supposedly given to us as rational. Accordingly, we cannot make sense of this supposedly given "knowledge" as genuinely knowledge. In other words, conceiving of knowledge of some aspect of reality as given to us, and intelligible only as such, undermines its very status as knowledge.

Stated in these terms, the problem with the Myth can seem obvious. However, many theories which attempt to explain our rational capacities can easily fall prey

to it. The problem arises when a theory aims to explain our rational capacities as depending on knowledge that can really be understood only as an achievement of those very capacities. That is the basic problem with the sense data theory that Sellars addresses in the beginning of *Empiricism and the Philosophy of Mind*. The sense data theorist wants to conceive of our capacity to for conceptual knowledge as dependence on our knowledge of sense data, but any knowledge of sense-data that we might have can really only be understood as an achievement of our capacity for conceptual. As such, insofar as this circularity is avoided, the sense data theorist left with a conception according to which our knowledge of sense data is intelligible to us only as given, and thus, not really intelligible as knowledge at all. So, generally, one is led into the Myth when one attempts to give an account of some aspect of our rational capacities as dependent on some sort of knowledge or awareness that, given the explanatory project for which it is recruited, must be conceived of as not involving an actualization of the rational capacities that it does in fact involve. The problematic structure of Givenness that arises in such a case is nicely stated by John McDowell (2009) as follows:

Givenness in the sense of the Myth would be an availability for cognition to subjects whose getting what is supposedly Given to them does not draw on capacities required for the sort of cognition in question.

If one is committed to such an explanatory structure, they face the following dilemma. On the one hand, they can refuse to acknowledge the contribution of rational capacities whose actualization is essentially involved in the knowledge to which they appeal, and thus are saddled with a view in which the sort of knowledge that they take to underlie the capacities they are trying to explain is *unintelligible*. On the other hand, they can acknowledge the capacities that are in fact essentially involved in the knowledge they appeal to in order to explain these very capacities, and thus are saddled with a view in which their account of these capacities is *incoherent*. This dilemma, which will arise repeatedly in various

forms throughout this dissertation, is the basic way in which the Myth manifests itself in a dialogical context.

Now, McDowell, here and in much of his other work, follows Kant in emphasizing that our rational capacities are essentially *conceptual* capacities. Sellars is a Kantian that has undergone a linguistic turn. The crucial point for Sellars is that our conceptual capacities are essentially *linguistic* capacities. As he puts it “grasping a concept is always mastering the use of a word.” Accordingly, if Sellars is right to follow Kant in thinking that awareness of anything of any cognitive significance requires the deployment of concepts, and the deployment of concepts is essentially a linguistic affair, then it follows that “all awareness of sorts, resemblances, facts, etc., in short, all awareness of abstract entities—indeed, all awareness even of particulars—is a linguistic affair” (1997, 63).¹¹ With this claim on board, it is not hard to see what the problem for worldly semantics, from Sellarsian perspective, is. The problem for worldly semantics is that semantic competence is supposed to be explained as depending on knowledge of the world. This explanatory structure presupposed here requires that knowledge of the world *does not* presuppose semantic competence. However, insofar as this worldly knowledge is a product of rational capacities, rational capacities are essentially conceptual capacities, and grasping a concept is always mastering the use of a word, this knowledge *does* presuppose semantic competence. To put the problem in terms of McDowell’s formulation of the problematic structure of Givenness, in articulating a worldly semantics for some language that some subjects speak, we require an availability for cognition of worldly knowledge to these subjects such that their getting this worldly knowledge does not draw on the capacities that are in fact required for having it, specifically, the capacity to

¹¹This is Sellars’s so-called “psychological nominalism.” Many have said that this is not a very good name for his position. I think otherwise, but I won’t get into such an argument here. For an explanation of the sense of Sellars’s term and its connection to the more familiar ontological sense of “nominalism,” see Simonelli (M.S.)

use and understand sentences of that language.

Now, as I said above, to limit the scope of my discussion here, I have focused my attention on one aspect of our semantic competence that is supposed to be explained by a worldly semantic theory: our knowledge of facts consisting in sentences standing in certain relations of entailment and incompatibility in virtue of meaning what they do. On a worldly semantics, this semantic knowledge is taken to be asymmetrically dependent on worldly knowledge of what we might speak of the “metaphysical structure” reality. Worldly semantic frameworks, of both the extra- and intra-worldly variety are ultimately committed to a view in which the metaphysical structure of reality, be it articulated in terms of facts consisting in set-theoretic relations obtaining between extra-worldly entities or primitive modal relations obtaining between intra-worldly entities, is simply given to a potential learner of a language, such that that learner can map words, phrases, and sentences with their semantic values. Accordingly, they preclude us from being able able to non-circularly comprehend the worldly knowledge which is supposed to underlie our knowledge of meaning as genuinely knowledge. Of course, so far, I have just stated this claim. I have not yet given any argument that worldly semantic theories are problematic in the way that I have claimed they are. That is what I will do in the next two chapters, arguing against the two most common incarnations of worldly semantics in the contemporary literature. Before I begin my attack on worldly semantics, however, I want to clarify my targets, especially “extra-worldly semantics,” at which I will take aim in the next chapter.

1.7 Elucidatory and Explanatory Models in Semantic

Extra-worldly semantics is by far the most common type of semantics practiced by contemporary semantic theorists. In the next chapter, we’ll see that, if taken to

constitute an explanation of what it is in which speakers' knowledge of meaning consists, it involves a clear instance of the Myth. Many semantic theorists who employ such a framework, however, phrase what they are doing in such a way so as to not commit themselves to the claim that they really are *explaining* the knowledge of meaning that speakers have. Rather, they often put things so as to suggest that they are doing something else: *elucidating* or *explicating* this knowledge of meaning. Rarely, however, are theorists explicit about what this distinction is or what falling on one side of it rather than the other amounts to.

Consider, for instance, what Chierchia and McConnell-Ginet (1990) say in their introductory semantics textbook when sensing possible trepidation from their scientifically-minded reader about the appeal to "possible worlds" in the semantic theory:

[U]sing the formal framework of possible worlds in semantics has produced some very interesting and nontrivial accounts of various intensional phenomena, and many quite enlightening semantic studies have been generated. It certainly seems to us a fruitful hypothesis that our semantic competence can be elucidated in this framework, (207-208).

Chierchia and McConnell-Ginet say here that it seems to them to be a fruitful hypothesis that our semantic competence can be *elucidated* in the framework of possible worlds. They do not say that it seems to them that our semantic competence can be *explained* with the use of such a framework. Tellingly, when they originally lay out what their aim, as linguists, is, at the beginning of the book, they do so in such a way as to suggest that they are uncomfortable about the use of the expression "explain" in this context

[A]s linguists, our focus is on modeling the cognitive systems whose operation in some sense "explains" linguistic phenomena.

They never say in *what* sense the operation of the cognitive systems they seek to model "explains" the linguistic phenomena with which they are concerned, and

they never say why they put the expression “explains” in scare-quotes here. As they proceed in the book, this issue gets lost entirely. They drop this guardedness about the use of the expression “explains,” and freely talk about the semantic theory “explaining” and “accounting for” empirical phenomena such as “judgments of semantic relatedness,” using these expressions more or less interchangeably (51). This wavering between an elucidatory and an explanatory conception of semantics makes it difficult to determine to what extent the theories put forward by Chierchia and McConnel-Ginnet are targets of the attack on extra-worldly semantics put forth in next chapter. I do not deny that our semantic competence can be *elucidated* with the use of a framework that centrally employs the notion of possible worlds. Indeed, I think it can be, and I think that this elucidation can indeed be enlightening. What I am denying is that our semantic competence can be *explained* or *accounted for* with the use of such a framework. This crucial distinction, I believe, is often lost in contemporary theorizing about meaning, and it is this tendency of contemporary theorists to lose this distinction that is largely responsible for the pervasiveness of the Myth in contemporary theorizing about meaning.

One way to get the distinction between semantic theories that aim at elucidation and those that aim at explanation into view is to consider whether a certain sort of circularity is acceptable in the theory. A circular explanation, in which the facts that are supposed to be explained are appealed to in order to arrive at the “explanation” of them, is no explanation at all. If a semantic theorist is able to rest happily while being aware of circularity in their theory, then it is a good bet that they take themselves to be doing elucidatory rather than explanatory work. One such theorist is, Jaako Hintikka (1975), who, when considering the question of whether possible worlds could only be understood by reference to counterfactual claims which would then be understood in terms of a possible worlds semantics, writes,

[A] circle of explication need not be a vicious one, provided it is wide enough to enable a logician to uncover nontrivial aspects of the structure of the concepts involved, (135).¹²

Here, Hintikka uses the term “explication” rather than the term “explanation,” and this, of course, is no accident. A circle of *explanation* can only be a vicious one, but a circle of *explication* or *elucidation* need not be vicious. Stina Bäckström provides a clear statement of how at least some elucidations might be virtuously rather than viciously circular:

There is [. . .] at least one kind of case in which a philosophical account can be circular without fault, and that is when the account aims at elucidating two concepts or phenomena that are mutually interdependent. In that case, circularity—far from being a deficiency—is a necessary feature of a successful account, (2016, 192).

If worldly knowledge and semantic knowledge are conceived of as mutually interdependent, then it is possible that they can be mutually elucidated by an account that appeals to one in explicating the other and vice versa. This is essentially how conceptions of possible worlds that define them with the use of “meaning postulates” (Carnap 1952, Partee 2005) suggest that we think of things (whether the proponents of such conceptions know that they are suggesting this or not). I will discuss such conceptions in more detail in the next chapter, but the point of bringing them here is just to show what a self-consciously elucidatory rather than explanatory conception of worldly semantics could be. On such a conception, our semantic knowledge is explicated as depending on our worldly knowledge, but this worldly knowledge is analyzed in terms of our knowledge of relations among sets of possible worlds, and possible worlds are defined as depending on our semantic knowledge, made explicit by the laying down of “meaning postulates.” Clearly, a possible worlds semantics that is structured in

¹²I should note, Hintikka himself does not “officially” make this claim. It actually occurs in the context of a fictional dialogue with Quine, but it is clear that it represents Hintikka’s own view.

this way will not be able to explain or account for the knowledge of meaning that we have, since this knowledge must be appealed to in order to define the materials with which the theory is constructed, but it may very well be able to uncover non-trivial structural features of the meanings of which we have knowledge through the mutual elucidation of our semantic knowledge and our worldly knowledge.

If elucidating structural features of our semantic knowledge is all you are aiming to do in employing a possible worlds semantics, and you recognize the space for a semantic theory that actually explains this semantic knowledge, then you and I have no quarrel. There are many ends other than explanatory ones with which a possible worlds semantics that elucidates the structure of the space of meanings of which we have knowledge can aid us. Even in the context of an explanatory project, an elucidatory semantic theory can function to get the exp-landum into view, and an explanatory theory can function to explain it. The great merit of an extra-worldly semantic framework is that it enables us to provide characterizations of semantic relations and semantic operations in set-theoretic terms. For instance, it enables us to provide a characterization of the semantic relations of entailment and incompatibility and semantic operations of conjunction and negation in terms of the set-theoretic relations of being a subset of and being disjoint and the set-theoretic operations of intersection and complementation. Once again, I don't deny that the structure of the semantic relations that complex sentences stand to one another can be elucidated set-theoretically by thinking of these sentences and their parts as having sets of possible worlds and related mathematical entities as their semantic values. What I deny is that to assign sets of possible worlds to sentences is to model their meaning in a way that enables us to explain that they stand in these semantic relations or our knowledge that they do.

Because the distinction between elucidation and explanation gets lost, many semanticists take themselves to be explaining our knowledge of meaning when

all they can really be doing is elucidating it, and this leads them to take there to be no room for semantic theories that really are of the sort to be able to explain it. For instance, Paul Portner (2005), in his introductory textbook to formal semantics, relevantly entitled *What Is Meaning?*, says that one of the main reasons for thinking of meaning in terms of a possible worlds semantics of the sort introduced here is that this sort of semantics “lets us define some basic semantic concepts: synonymy, contraiety, entailment, contradiction, tautology,” (18). Portner makes this claim in the course of arguing that possible worlds semantics, as opposed to the holist empiricist semantics proposed by W.V. Quine (1953, 1960) or the social-normative semantics proposed by Robert Brandom (1994, 2000), is the sort of semantic theory with we should think about what meaning is (Portner 2005, 4-22).¹³ Possible worlds semantics, Portner claims on behalf of mainstream formal semantics, gives us a better account of what meaning is than the sort of semantic theories proposed by Quine or Brandom. If all possible worlds semantics is doing, however, is elucidating and not explaining, then possible worlds semantics does not give us a better account of what meaning is than the sorts of theories proposed by Quine or Brandom because it gives us no account of meaning at all.

By the end of this dissertation, I will have articulated a semantic theory that can actually account for what meaning is and what it is for us to grasp the meanings that we do. In contrast to the worldly semantic theories, on the theory I’ll propose, knowledge of meaning is not undergirded by worldly knowledge. On the contrary, our “worldly” knowledge really is *nothing other than* our knowledge of meaning, expressed in a worldly mode. On this Sellarsian story, the “worldly knowledge” to which the worldly semanticist appeals, is conceived of as a “shadow” of our knowledge of meaning which is, in reality, not worldly, but normative. I will then provide a normative semantics where the behavioral

¹³It is worth noting that, while Portner drops a footnote to Quine’s “Two Dogmas” and *Word and Object*, it is hard to see how the view he characterizes as Quine’s really is Quine’s. He seems to be failing to discriminate Quine’s empiricist holism from Brandom’s rationalist holism.

patterns codified by these norms can be explained without appeal to knowledge of the worldly entities that the worldly semanticist takes to be contained in meanings. This will make space for an account of the real relation between language and the world, which, once again drawing from Sellars, I will provide. All of that in good time. But first things first: I must argue that worldly semantics, of both the extra- and intra-worldly variety, at least insofar as the semantic theories have explanatory rather than merely elucidatory aims, really do contain, at their very core, instances of the Myth of the Given.