BEYOND RIGIDITY

THE UNFINISHED SEMANTIC
AGENDA OF
NAMING AND NECESSITY

SCOTT SOAMES

BEYOND RIGIDITY

THE UNFINISHED SEMANTIC AGENDA OF
NAMING AND NECESSITY

SCOTT SOAMES

OXFORD
UNIVERSITY PRESS

2002

Preface

In *Naming and Necessity*,[1] Saul Kripke argued that proper names and natural kind terms are rigid designators, and that, in part because of this, their meanings are not given by the descriptions that speakers associate with them. In so doing, he told us what the meanings of these expressions are not; however, he did not provide a positive account of what their meanings are. I do so in this book.

In the first part of the book (chapters 1–8), I argue that for a great many proper names, meaning and reference coincide. This view is, of course, not new, having been championed by John Stuart Mill, as well as by a number of contemporary neo-Russellians, most notably Nathan Salmon.[2] I myself have long been a defender of the Millian view that the meanings of most linguistically simple proper names are their referents, and of the Neo-Russellian view that the proposition semantically expressed by an attitude ascription containing such a name in its content clause reports a relation between an agent and a singular, Russellian proposition. In this book I ground these views in a larger, explanatory conception of meaning, and of semantic content, together with an accompanying account of how the semantic content of a sentence relates to information conveyed and asserted by utterances of the sentence in different contexts. A central feature of this account is the explanation it provides of how sentences containing names or indexicals may be used to convey, and even assert, propositions the contents of which exceed the semantic contents of the sentences uttered.

In chapter 1, I describe two main items of the unfinished semantic agenda of *Naming and Necessity*: the development of a positive theory of the meaning, and semantic content, of proper names, and the proper extension of the central semantic theses about names to the more linguistically diverse, and philosophically significant, class of natural kind terms. In chapter 2, I show how Kripke's argument that proper names

are rigid designators, and therefore are not synonymous with nonrigid descriptions associated with them by speakers, can be extended to rule out the possibility that names are rigidified descriptions, or descriptions that are required to take wide scope in modal contexts. In chapter 3, I develop an account of meaning that justifies the claim that the meanings of many, but not all, proper names are their referents. In chapter 4, this account is refined in order to account for the ambiguity of many proper names, and it is generalized to include the semantic contents of indexicals. In chapter 5, I discuss an important class of exceptions to the theses about proper names developed in earlier chapters. Here I discuss a rich and varied class of linguistically complex phrases that I call *partially descriptive names*. These names, though not, strictly speaking, rigid designators, are nearly so, in the sense that they always designate the same object, when they designate anything at all. In addition, many of them are standardly recognized as names, and they can be shown to display most of the important features of linguistically simple proper names. I argue that partially descriptive names are equivalent to certain quite special definite descriptions that combine direct reference to their referents with partial descriptions of them; the semantic contents of these expressions are amalgams of their referents plus additional descriptive information carried by the phrases as a whole. Chapters 6–8 are concerned with propositional attitude ascriptions, particularly those containing names or indexicals in their content clauses. I argue that it is possible to reconcile the combination of (i) a Millian account of the semantic contents of simple names and indexicals and (ii) a Russellian treatment of attitude ascriptions with (iii) Fregean intuitions about the information conveyed, and even asserted, by utterances of both simple sentences and attitude ascriptions containing such expressions.

The remainder of the book is concerned with the second main piece of unfinished business left to us by *Naming and Necessity*—the task of applying the semantic insights gained from the study of proper names to natural kind terms. Although I have long been persuaded of the basic correctness of the anti-descriptivist account of these terms developed by Saul Kripke and Hilary Putnam, I have also been puzzled by central aspects of it.[3] Kripke, in particular, models his treatment of natural kind terms on his account of proper names. For example, he characterizes both as rigid designators, and claims that, because of this, identity sentences involving such terms, including sentences expressing

theoretical identifications, are necessary if true. Over the years, this characterization of natural kind terms has come to be regarded by many as axiomatic of the Kripke-Putnam view. This is puzzling, since the only definition of rigid designation that Kripke ever gives applies solely to singular terms, whereas natural kind terms come in a variety of syntactic and semantic types—including mass nouns, count nouns, and adjectives, all functioning as predicates. This raises the question of what it might mean to characterize such a predicate as rigid designator.

This question is addressed in chapter 9, where I argue that there is no natural way of extending the concept of rigidity from singular terms to predicates that vindicates the central doctrines of *Naming and Necessity*. In particular, I argue that there is no natural concept of rigidity applicable to predicates according to which (i) all natural kind predicates are rigid, whereas familiar descriptive predicates, like *is a bachelor*, are not, and (ii) theoretical identification sentences involving rigid predicates must be necessary, if true. If this is right, then we need to find some other way of characterizing the semantic similarities between natural kind predicates and proper names. I do this in chapters 10 and 11, where I provide a different, more limited vindication of the claim that the semantics of natural kind terms guarantee that certain theoretical identification sentences involving them are necessary, if true. According to the view I develop, rigidity is not the key to the semantics of natural kind terms in general; nor is it central to bringing out the most important properties they share with proper names.

The views presented in this book have been in the works for a number of years, during which time I have benefited greatly from the contributions of many people and several institutions. The idea for the book originated in a series of lectures, "Logic in Natural Language," that I gave at the Lingua 98 conference, held in January 1998 at the department of informatics of the Federal University of Pernambuco in Recife, Brazil. In "Reference, Intentionality, and the Aims of Semantics," presented to a group of cognitive scientists, I tried to trace certain technical disputes in semantics—for example, disputes about propositional attitude ascriptions containing names or indexicals—to foundational questions about the nature of linguistic meaning, and its relation to information conveyed and asserted by utterances of declarative sentences. It was at this time that I developed the ideas behind both the approach to linguistic meaning and communication elaborated in chapter 3 and the

application of that approach to attitude ascriptions containing proper names given in chapter 8. Later versions of this material provided the basis for lectures at the University of California at Santa Barbara in April 1999, at the Center for the Study of Language and Information at Stanford University in May 1999, and at the conference on Methods in Philosophy and the Sciences at the New School for Social Research in December 1999.

My first presentation of material from chapters 9 and 10 was at an international conference on the work of Saul Kripke held at the Instituto de Investigaciones Filosóficas, Universidad Nacional Autónoma de México, in Mexico City in October 1996. Later versions of this material were presented in lectures at the Graduate Center of the City University of New York in March 1997, at Harvard University in April 1997, at UCLA in November 1997, at Pernambuco in January 1998, at Cornell University in April 1998, at the University of California at Davis in November 1998, at Stanford University in March 1999, at Arizona State University in March 1999, at California State University at Northridge in May 1999, and at Ohio University in October 1999.

In addition, an early version of the book manuscript as a whole (minus chapter 5) was presented during a yearlong graduate seminar at Princeton that I taught jointly with David Lewis during the 1999–2000 academic year. I am much indebted to David and other participants in that seminar, including Mark Johnston, Kit Fine, Jonathan Vogel, Cian Dorr, Benj Hellie, Michael Nelson, and Jonathan McKeown-Green for extremely valuable input that helped shape the final formulations of many of my views.

Although the main ideas presented in this book have not appeared in print before, parts of chapters 2 and 7 contain material originally published elsewhere. With the exception of its final section, which is new, chapter 2 is an updated and expanded version of "The Modal Argument: Wide Scope and Rigidified Descriptions," *Nous*, vol. 32, 1998, 1–22. In addition, one section of chapter 7 includes some material that originally appeared in "Beyond Singular Propositions?," *The Canadian Journal of Philosophy*, vol. 24, 1995, 515–550.

Finally, I would like to express my gratitude to those who have made systematic contributions to this work. These include the philosophers Ali Kazmi, Michael Thau, James Pryor, and Jeff King, all of whom read early versions of the manuscript and provided numerous comments that resulted in substantial contributions to the final product.

The same is true of Kent Bach and the members of the Bay Area Philosophy of Language Discussion Group, which devoted three sessions in the spring of 1999 to parts of the book. My student Jeff Speaks proofread the penultimate version of the manuscript and made many helpful suggestions for final revisions. I would also like to express an intellectual debt to my friend and former colleague Saul Kripke, whose seminal contributions to the field both inspired and provided the theoretical framework for the present work.

Finally, special thanks are owed to the Center for Advanced Study in the Behavioral Sciences in Stanford, California, where I wrote the first draft of the book during the 1998–1999 academic year. My year there was financed in part by Princeton University, from which I was on sabbatical, and in part by the Andrew W. Mellon Foundation. I am grateful to these institutions for providing me the time, free of competing distractions, to focus on this work.

Princeton, New Jersey   Scott Soames
Spring 2001

Contents

BEYOND RIGIDITY

1

The Unfinished Semantic Agenda
of *Naming and Necessity*

This book is concerned with rigid designation in particular and, more generally, with the unfinished semantic agenda that has been left to us by Saul Kripke's *Naming and Necessity*. Kripke's strategy in *Naming and Necessity* is to begin by articulating semantic doctrines covering the simplest case, proper names, and then to extend his theory to the more complex and potentially significant case of natural kind terms. Along the way, questions about the modal and epistemic status of sentences containing proper names and natural kind terms come in for extended discussion. Modal considerations give rise to doctrines about the truth conditions of sentences in a rich and substantial sense— namely, as conditions that possible states of the world must satisfy if sentences are to be true when taken as descriptions of those states. Epistemic considerations raise fundamental questions about the semantics of attitude ascriptions like *x knows/believes/asserts that S*.

These modal and epistemic considerations take us to the heart of the semantic enterprise. The simplest and most fundamental question to be answered by a semantic theory is *What do sentences say or express (relative to various contexts of utterance)*? This, in turn, is closely related to the question *What do speakers say, and what beliefs do they express, when they assertively utter sentences of their language*? A semantic theory that tells us what sentences say helps us answer this latter interpretive question about speakers, in virtue of principles like (1).

1. A sincere, reflective, competent speaker who assertively utters S in a context C typically asserts (among other things) what S says in C.

This principle presupposes a relational analysis of the attitude of saying or asserting—an analysis that sees it as a relation between speakers, who do the asserting, and the semantic contents of sentences, which are the things asserted. Once this analysis is accepted, it is natural to view propositional attitude reports in accord with (2) and (3).

2. An individual i satisfies *x says (asserts) that S*[1] relative to a context C iff i stands in a certain relation R, the assertion relation, to the semantic content of S in C.
3. An individual i satisfies *x v's that S* (where v = 'believes,' 'knows,' 'proves,' etc.) iff i stands in a certain relation R' to the semantic content of S in C.

I will call the semantic content of a sentence relative to a context, the proposition semantically expressed by the sentence relative to that context. My working hypotheses will be (i) that the central task of a semantic theory is to specify a function from sentence-context pairs to propositions semantically expressed by the sentence in those contexts; (ii) that propositions are the objects of propositional attitudes in the sense of principles (2) and (3); and (iii) that a sentence, when set in a context C, is true with respect to an arbitrary possible world-state w iff the proposition expressed by S in C is true with respect to (i.e., when taken as a description of) w. According to this framework, the semantic analysis of an expression is a theory about the propositions expressed by sentences containing the expression. This theory has immediate consequences for the truth conditions of sentences, including propositional attitude ascriptions, containing the expression. This is the perspective from which I investigate proper names, natural kind terms, and related expressions.

Let us begin with a summary of central theses about proper names defended in the first two lectures of *Naming and Necessity*.

Theses About Proper Names

T1. Proper names are rigid designators: a proper name that designates an object o does so with respect to all world-states in which o exists, and never designates anything else.
T2. Proper names are nondescriptional: (i) they are not synonymous with descriptions or clusters of descriptions associated

with them by speakers; (ii) the referent of a name with respect to an arbitrary world-state w is not determined semantically via the satisfaction of any description or descriptive condition at w; instead, (iii) the referent of a name is initially fixed at the actual world-state and, once fixed, is stipulated to remain the same with respect to all other world-states.[2]

T3. The referent of a proper name is initially determined in one or the other of two ways: by an ostensive baptism or by a stipulation that it is to be whatever satisfies a certain description. Later, when the name is passed from speaker to speaker, the way in which the reference was initially established usually doesn't matter. Typically, speakers farther down the historical chain use the name to refer to the initial referent whether or not they associate properties with the name that (uniquely) apply to it.
T4. Identity sentences in which different names (or other rigid designators) flank the identity sign are necessary if true. Nevertheless, often the truths expressed by these sentences are knowable only a posteriori.

In chapter 2 I will look closely at the arguments behind theses T1 and T2, and examine certain descriptivist counterclaims—that names are equivalent to rigidified descriptions, or that names are equivalent to descriptions that are required to take wide scope over modal operators. For the moment, however, let us assume that the arguments for T1 and T2 go through, and that these theses are correct. If they are correct, then presumably the semantic content of a proper name is not the same as that of any description, and the proposition semantically expressed by a sentence containing a name is not the same as the proposition expressed by any corresponding sentence in which a description has been substituted for the name.

This is an interesting negative result. However, it is not accompanied by any corresponding positive result. Nowhere in *Naming and Necessity*, or anywhere else, does Kripke tell us what the semantic content of a name is; nor does he tell us precisely what proposition is expressed by a sentence containing a name. The perplexing nature of this gap in his analysis may be brought out by the following speculation: If the semantic content of a name is never the same as that of any description, then it seems reasonable to suppose that names don't have

descriptive senses, or descriptive semantic contents, at all. Moreover, if names don't have descriptive semantic contents, then it would seem that their only semantic contents are their referents. From this it follows that coreferential names have the same content. If we add a plausible principle of compositionality, we are led to the view that sentences differing only in the substitution of one of those names for another must have the same semantic content, and so must semantically express the same proposition. However, this conclusion plays havoc with thesis T4. For, on this line of reasoning, if a and b are proper names, and the sentence $a = b$ is true, then it semantically expresses the same proposition as the sentence $a = a$. But then, since the proposition expressed by $a = a$ is surely knowable a priori, so is the proposition expressed by $a = b$.[3]

This conclusion conflicts with T4. We can therefore be sure that, at least at the time of *Naming and Necessity*, Kripke didn't accept it. What is not completely clear is why he didn't. The alleged necessary aposterioricity of the truths expressed by identity sentences in which different names flank the identity sign is discussed at some length at the end of lecture 2 of that work.[4] The view Kripke presents there goes essentially as follows: Let $a = b$ be a true identity sentence involving proper names. These names may either be ordinary names like *Cicero* and *Tully*, or names like *Hesperus* and *Phosphorus*—where understanding the latter may involve associating them with specific reference-fixing descriptions. Either way, Kripke argues, the evidence available to a competent user of the names—just by virtue of understanding them—is insufficient to determine that the names are coreferential. He illustrates this by noting that there is a possible state of the world in which speakers are in an evidentiary situation qualitatively identical with the one in which we actual speakers find ourselves, and yet in the merely possible situation the names are used to refer to different things. For example, there is a possible state of the world in which speakers fix the referent of the name *Hesperus* just as we do in the actual world—by pointing to a bright object that appears in the evening in a certain part of the sky in certain seasons. Furthermore, speakers in that possible state fix the referent of the name *Phosphorus* by pointing to a bright object that appears in the morning in certain seasons. From a qualitative point of view, these speakers are in the same evidentiary situation with respect to their uses of the names as we are. Yet in their state the names are used to refer to different things.

Kripke intends this example to show that the evidence available to agents in any of these possible states of the world, simply by virtue of being competent users of the names, is insufficient to show that the names are coreferential. We may express this idea as follows: Let E be the collection of possible world-states in which the epistemic situation of agents regarding their uses of the terms *Hesperus* and *Phosphorus* is qualitatively identical with our actual epistemic situation. One might then think that any proposition which fails to be true in all members of E is a proposition which is not determined to be true by the qualitative evidence available to us, and so is one that we do not know a priori, simply on the basis of our mastery of the relevant terms or concepts. Let us suppose this is right. Well, one proposition that fails to be true in all members of E is the proposition that the names *Hesperus* and *Phosphorus* are coreferential in our language; another closely related proposition is the proposition that the identity sentence *Hesperus = Phosphorus* expresses a truth in our language. Thus, we are in a position to conclude that the metalinguistic claim that the sentence *Hesperus = Phosphorus* is true in our language is something that is not knowable a priori; rather, it is something that we can come to know only on the basis of empirical investigation.

So far so good. However, there is a problem. The lesson Kripke explicitly draws from the example is not that a certain metalinguistic claim is knowable only a posteriori but, rather, that the claim that Hesperus is Phosphorus is knowable only a posteriori. This can be seen from the following passage in which he sums up his argument.

> The evidence I have before I know that Hesperus is Phosphorus is that I see a certain star or a certain heavenly body in the evening and call it 'Hesperus', and in the morning and call it 'Phosphorus'. I know these things. There certainly is a possible world in which a man should have seen a certain star at a certain position in the evening and called it 'Hesperus' and a certain star in the morning and called it 'Phosphorus'; and should have concluded—should have found out by empirical investigation—that he names two different stars, or two different heavenly bodies. At least one of these stars or heavenly bodies was not Phosphorus, otherwise it couldn't have come out that way. But that's true. And so it's true that given the evidence that someone has antecedent to his empirical investigation, he can be placed in a sense in exactly the same situation, that is a qualitatively identical epistemic situation, and call two heavenly bodies 'Hesperus' and 'Phosphorus', without their being identical. . . . So two things are true: first,

that we do not know apriori that Hesperus is Phosphorus, and are in no position to find out the answer except empirically. Second, this is so because we could have evidence qualitatively indistinguishable from the evidence we have and determine the reference of the two names by the positions of two planets in the sky, without the planets being the same.[5]

The problem with this passage is that Kripke's conclusion does not follow from his argument as stated. The proposition that Hesperus is Phosphorus is, as Kripke rightly insists, true in all possible states of the world. So it is true in all members of the class of world-states E in which agents are in an epistemic situation qualitatively identical to ours. And since it is true in those worlds, the principle that only propositions true in all members of E are known a priori does not rule out that it may be knowable a priori.

The point I am making depends on sharply distinguishing between (4a) and (4b).

4a.   Hesperus = Phosphorus
4b.   'Hesperus = Phosphorus' expresses a truth in our language.

When explaining the necessity of (4a), Kripke uses his example of the possible world-state whose agents are in an epistemic situation qualitatively identical to ours to remind us that the contingency of (4b) is irrelevant to the necessity of (4a). According to Kripke, the agents in his imagined world-state use the sentence *Hesperus = Phosphorus* to express a proposition different from the proposition we actually use it to express. The fact that the proposition they use it to express is false in their world-state does not show that the proposition we actually use it to express is false when evaluated in their world-state, or any other.

What Kripke fails to point out is that the same reasoning applies to the epistemic status of the two examples. Proposition (4b) is knowable only a posteriori. But that has no obvious bearing on the question of whether proposition (4a) is a priori. The agents of Kripke's imagined world do not know the proposition they use the sentence *Hesperus = Phosphorus* to express, for the simple reason that the proposition they use the sentence to express is false in their world. But this does not show that the different proposition we use the sentence to express isn't known by us; nor does it show that it isn't known by us independent of empirical investigation. For this reason, Kripke's conclusion—that

it is not knowable a priori that Hesperus is Phosphorus—does not follow from the considerations he adduces.

More precisely, it does not follow from these considerations alone. Perhaps if Kripke's explicit remarks were supplemented with some further principles, the gap in the argument could be filled. What is needed are principles connecting an agent's understanding and accepting a sentence, on the one hand, with the agent's believing or knowing the proposition expressed by the sentence, on the other.[6] One principle of this sort is the following.[7]

Strong Disquotation

A sincere, reflective, rational individual i who understands S is disposed to accept S iff i believes the proposition expressed by S, and thereby satisfies *x believes that S*.

Agents in an epistemic situation qualitatively similar to our situation before the astronomical discovery understand, but are not disposed to accept, *Hesperus is Phosphorus*, and so they don't believe what they express by the sentence. Similarly, prior to the astronomical discovery we didn't accept the sentence, so at that time we didn't believe that Hesperus is Phosphorus. Moreover, the evidence available to both of us by virtue of our understanding the terms is such that we would not have been **justified** in accepting the identity sentence on the basis of that evidence. With this in mind, one might formulate the following principle involving disquotation and justification:

Strong Disquotation and Justification

A sincere, reflective, rational individual i who understands S and is in possession of evidence e would be justified in accepting S on the basis of e iff i's possession of e, and i's reasoning correctly about it, would be enough to ensure that i would be justified in believing the proposition expressed by S, and hence that i satisfies *x would be justified in believing that S*.

If these two principles are accepted, then Kripke's argument can be reconstructed as follows:

(i)   Since there are possible situations in which *Hesperus is Phosphorus* expresses something false, even though the

agents in those situations are perfect reasoners who have evidence qualitatively identical with the evidence available to us simply on the basis of our linguistic competence, the evidence available to us simply on the basis of our linguistic competence does not justify our accepting the sentence.

(ii) So, by the strong disquotation and justification principle, the evidence available to us simply by virtue of our competence, plus our reasoning correctly about it, is not enough to justify us in believing that Hesperus is Phosphorus.

(iii) If the belief that Hesperus is Phosphorus were justifiable a priori, then it would be justifiable by virtue of the evidence available to us by virtue of our linguistic competence, plus our reasoning correctly about it.

(iv) Thus that belief is not justifiable a priori. Hence, it is not knowable a priori that Hesperus is Phosphorus.

Although this argument fills the gap in the passage from Kripke, it is not transparently sound. One potential problem is that step (i) seems to rely on a questionable general principle—namely, that if an agent could have evidence qualitatively identical with my evidence for accepting a certain sentence S, even though S is false in that agent's situation, then I would not be justified in accepting S on the basis of the evidence I possess. But consider my qualitatively identical twin in a merely possible world who lives a life identical with mine up until last night, at which time his brain is removed, placed in a vat, and artificially stimulated so as to have experiences qualitatively identical with my actual experiences. If today my twin were to accept *I am not a brain in a vat*, he would be accepting something false. If the general principle implicit in (i) were correct, this would mean that I am not justified in accepting *I am not a brain in a vat*. It is, however, far from obvious that I am not justified.

Second, and even more significant, the strong disquotational principles on which (ii) depends are troublesome in their own right. The central difficulty is illustrated by the case of puzzling Pierre, presented by Kripke in "A Puzzle About Belief," which appeared in 1979, nine years after Kripke delivered the lectures which became *Naming and Necessity*.[8] Kripke's Pierre is a Frenchman who grows up in Paris, speaking French. He sees picture postcards of London and forms the belief that London is pretty, which he expresses by saying *Londres est*

jolie. Later, he moves to London, learns English not by translation but by the immersion method, and lives in an unattractive part of the city. On the basis of his experience he forms a belief that he expresses by saying *London is not pretty*. It is not that he has given up the belief he formed in Paris on the basis of the picture postcards. He still affirms *Londres est jolie* when speaking French to old friends, even though he does not accept the English sentence *London is pretty*. The reason for this disparity is that he doesn't realize that *Londres* and *London* are names of the same city. This doesn't mean that he fails to understand the two sentences. He understands the former as well as he and his French-speaking friends did while he was living in France, which was certainly well enough to assert and communicate his belief that London is pretty, and he understands the latter as well as monolingual native speakers of English, who surely count as competent speakers, do. Moreover, since the sentences are translations of one another, they express the same proposition. But now we have a problem. By the strong disquotation principle (right-to-left direction) we get the result that Pierre does not believe that London is pretty, because he understands but does not accept *London is pretty*. By a corresponding strong disquotational principle for French (left-to-right direction), together with an appeal to a standard translation from French to English, we get the result that Pierre does believe that London is pretty. Since this is a contradiction, we have an apparent reductio ad absurdum of the strong disquotational principles.[9]

One way of putting the general idea behind the strong disquotation principles is that in order to believe a proposition, one must be disposed to accept every sentence one understands that expresses that proposition. As Kripke's Pierre example illustrates, this idea overlooks the possibility that an individual might understand two sentences that express the same proposition, without knowing that they do, and so might accept one of the sentences while not accepting the other.[10]

Another illustration of this possibility is provided by the following example, due to Nathan Salmon.[11] Salmon asks us to imagine an individual, Sasha, who learns the words *catsup* and *ketchup* by independent ostensive definitions, perhaps by being given, at different times, bottles with these words on the labels to season his foods. As a result of these experiences Sasha comes to learn what catsup is and what ketchup is. However, since the occasion never presents itself, no one ever tells him that the two words are synonymous (which of course they are). Because

of this, Sasha does not accept the identity sentence *Catsup is ketchup*— either because he suspects that there may be some, to him indiscernible, difference between the two, or because he thinks it improbable that different words would be used for the same condiment. Though this is unusual, it does not disqualify him from understanding the two terms. As Salmon emphasizes, nearly all of us learn one of the terms ostensively, before learning the other. The order in which the terms are learned doesn't matter, and if either term may be learned in this way, then surely it is possible that someone like Sasha could learn both ostensively, without being told that they are synonymous. But if this is right, then there will be sentences S_c and S_k which differ from one another only in the substitution of one term for the other such that Sasha understands both while being disposed to accept only one, say S_c. But then by strong disquotation we get the result that *Sasha believes that S_c* is true whereas *Sasha believes that S_k* is not true—which is impossible, given the synonymy of the two sentences.

Indexicals provide further problems of a similar sort. For example, imagine the following case: Professor McX, dazzled by the performance of his school's quarterback, points at the player and says, "He is the finest athlete in school." The quarterback's friend overhears the remark, and when he sees him in their next class, tells him, "Our math teacher, Professor McX, believes that you are the finest athlete in school." The friend's remark is, of course, true; McX **does** believe this. However, McX, who overhears the remark while not recognizing his student as the player he saw, denies it. Speaking to the quarterback, McX says, "You, the finest athlete in school? Don't be silly. You are a math genius, and they are never good athletes." Since the professor is not disposed to accept *You are the finest athlete in school* in this context (in which the student/quarterback is the referent of *you*), the context-relativized principle of strong disquotation (needed for sentences containing indexicals)

Context-Relativized Strong Disquotation

A sincere, reflective, rational individual i who understands S is disposed to accept S in a context C iff i believes the proposition expressed by S in C, and thereby satisfies *x believes that S* in contexts that incorporate the same world-state and assign the same semantic values to context-sensitive expressions as C.

gives the incorrect result that McX does not believe the proposition expressed by *You are the finest athlete in school* in contexts in which the student/quarterback is the addressee.

The upshot of all this is that there are serious doubts about the strong disquotational principles. We need not, at this point, regard these principles as definitely refuted, but neither should we accept them.[12] However, if we don't accept them, then the gap in Kripke's argument that the statement *Hesperus is Phosphorus* is an example of the necessary a posteriori remains unfilled. More generally, his discussion of thesis T4 in *Naming and Necessity* provides no compelling argument for the claim that identity statements made using sentences in which different but coreferential proper names flank the identity sign are standardly knowable only a posteriori.[13]

How bad is this? Many, I think, take it simply to be a datum that one can know that Cicero is Cicero, or that Hesperus is Hesperus, without knowing that Cicero is Tully, or that Hesperus is Phosphorus. In the presence of my proposed relational analysis of propositional attitudes and attitude ascriptions, this alleged datum leads to the conclusion that sentences which differ only in the substitution of coreferential proper names may semantically express different propositions. And if the sentences $a = a$ and $a = b$ express different propositions, then at least the observation that the proposition expressed by the former is knowable a priori won't force the conclusion that the proposition expressed by the latter must also be knowable a priori. Thus, even though Kripke hasn't given us a persuasive argument that the proposition expressed by $a = b$ is knowable only a posteriori, accepting the alleged datum would at least allow one to block the most obvious line of argument to the effect that it must be knowable a priori.

But now there is a difficulty to be faced. We need some positive account of the contributions made by proper names to the propositions semantically expressed by sentences containing them. Moreover, if the alleged datum is to be accepted, then this account must make clear precisely in what respect the propositions semantically expressed by sentences containing different but coreferential proper names differ. What makes this task so daunting is that the old solution to this problem—the view that names have descriptive semantic content—seems to have been thoroughly discredited by the arguments of Kripke and others. If this is right—if the idea that names have descriptive semantic content really has been discredited—then, given the alleged datum, one

cannot identify the semantic contents of names either with their referents or with descriptive information that may vary from one coreferential name to another.

It is not clear what alternatives remain. In what other ways do coreferential names differ? Often, they have different syntactic, phonological, and orthographic properties, and in theory one could appeal to these differences to distinguish the different propositions semantically expressed by sentences containing different names. Surely, however, we don't want to say that speakers using names that differ syntactically, phonologically, or orthographically can never assert or believe the same proposition.

Thus we are left with a dilemma. On the one hand, we may accept the alleged datum that typically, when sentences differ only in the substitution of one proper name for another, it is possible to assert and believe the proposition semantically expressed by one of the sentences without asserting and believing the proposition semantically expressed by the other. If we do this, then we must give some positive account of propositions and propositional attitudes that explains how this is possible. This problem is exceedingly difficult if, as Kripke seems to have shown, proper names do not have descriptive semantic contents. On the other hand, we may reject the alleged datum and identify the semantic contents of names with their referents. If we do this, we will be led to maintain that sentences which differ only in the substitution of coreferential names semantically express the same propositions, and that attitude ascriptions involving such sentences are truth-conditionally equivalent. The difficulty for this approach is to explain how, if it is correct, speakers succeed in using such sentences to convey different information and express different beliefs, which they clearly do, and why speakers often do not regard attitude ascriptions involving such sentences to be truth-conditionally equivalent, which they frequently do not. Dealing with this dilemma is one of the two most important pieces of unfinished semantic business left to us by *Naming and Necessity*. A central part of my task in chapters 2–8 will be to resolve it.

The second major unresolved semantic problem that we have inherited from *Naming and Necessity* is an extension of the first. Until now I have been concentrating on proper names, reviewing Kripke's theses T1–T4, and stressing the important questions left unanswered by them. At this point, it is good to remind ourselves that the semantic model that Kripke presented was never intended to be confined to

proper names alone. On the contrary, much of the significance of his work is owing to the fact that he intended to include the rich and heterogeneous class of natural kind terms in its scope. Unfortunately, this aspect of his semantic model was never fully specified. As a result, it has remained incompletely developed and poorly understood. Thus, the second important piece of unfinished semantic business left to us by *Naming and Necessity* is that of understanding how to extend the model of proper names to other classes of expressions, including natural kind terms of different grammatical categories.

I have already summarized the four central theses about proper names defended in the first two lectures of *Naming and Necessity*. In lecture 3, similar theses are defended for natural kind terms. For example, Kripke argues at length that natural kind terms like *gold, tiger, cat, water, heat,* and *light* are not synonymous with clusters of descriptions standardly associated with them by speakers. As in the case of proper names, two ways are given by which the reference of a term may initially be fixed. One way involves direct presentation of samples of the putative kind, together with the stipulation that the term is to be understood as applying to all and only instances of the unique natural kind (of a certain sort) of which nearly all members of the sample are instances. The other way of fixing the reference of a natural kind term involves the use of a description that picks out the kind, or members of the kind, by some, usually contingent, properties. Later, when the kind term is passed from speaker to speaker, the way in which the reference was initially fixed normally doesn't matter—just as with proper names.[14] As a result, speakers farther down the linguistic chain may use the term to apply to instances of the given kind, whether or not the descriptive properties they associate with the term really pick out members of that kind.

In addition, scientific investigation may lead to the discovery of properties that are necessary and sufficient for membership in the kind. These properties are expressed in theoretical identity sentences that express truths that are necessary but a posteriori. Examples of such sentences specifically discussed in *Naming and Necessity* are *Water is H_2O* (pp. 126–129), *Flashes of lightning are flashes of electricity* (p. 132), *Light is a stream of photons* (pp. 129–130), *Gold is the element with atomic number 79* (pp. 123–125), *Cats are animals* (pp. 122–123), *Whales are mammals* (p. 138), *Heat is the motion of molecules* (pp. 99–100).

The parallels between Kripke's treatment of proper names and his discussion of natural kind terms are evident. However, there are special complications that arise in the discussion of natural kind terms. Among the most important of these are questions about rigidity, and related questions about the modal properties of certain identity sentences. As in the case of proper names, natural kind terms are said to be rigid, and the putative rigidity of these terms is used to support the corollary that theoretical identity sentences involving them are necessary, if true. For example, in discussing theoretical identifications involving natural kind terms, Kripke says *"Theoretical identities, according to the conception I advocate, are generally identities involving two **rigid** designators and therefore are examples of the necessary aposteriori."*[15] Another example occurs in the discussion of the mind/body identity theory, where he maintains that *pain* and *c-fiber* stimulation are rigid designators, and adds *"So it seems that the identity theorist is in some trouble, for, **since we have two rigid designators, the identity statement in question is necessary.**"*[16] Again, this remark carries the implication that any identity sentence involving rigid designators will be necessary, if true.

However, there is a difficulty here that has not been widely appreciated. Kripke gives no separate definition of what it means for a natural kind term to be rigid; nor does he provide distinct arguments to show that such terms are rigid. This is a problem because his explicit definition of rigidity tells us only what it is for a singular term to be rigid.[17] If all natural kind terms were just ordinary singular terms, each purporting to designate a single object, then this definition could be applied directly to them, without qualification. However, as Kripke recognizes, natural kind terms fall into a variety of syntactic and semantic categories.

For example, he says, *"According to the view I advocate, then, terms for natural kinds are much closer to proper names than is ordinarily supposed. The old term 'common name' is thus quite appropriate for **predicates** marking out species or natural kinds, such as 'cow' or 'tiger.' My considerations apply also, however, to certain mass terms for natural kinds, such as 'gold', 'water' and the like."*[18] A little later, summing up his views, Kripke adds:

> . . . *my argument implicitly concludes that certain general terms, those for natural kinds, have a greater kinship with proper names than is generally realized. This conclusion holds for certain for various species names,*

> *whether they are count nouns, such as 'cat,' 'tiger,' 'chunk of gold,' or mass terms such as 'gold,' 'water,' 'iron pyrites.' It also applies to certain terms for natural phenomena, such as 'heat,' 'light,' 'sound,' 'lightning,' and, presumably, suitably elaborated, to corresponding adjectives—'hot,' 'loud,' 'red.'*[19]

It appears from these passages that Kripke intends his general theses about natural kind terms—including, presumably, the claim that they are rigid—to apply in some form to terms of various syntactic and semantic categories. This raises a number of fundamental questions. Included among the questions with which I will be concerned are the following:

(i) What is it for a predicate to be a rigid designator?

(ii) Are natural kind predicates, like *cow, tiger, animal, chunk of gold, flash of lightning*, and *drop of water* rigid?

(iii) What sorts of sentences count as identity sentences involving predicates? What are their logical forms? In particular, can they be taken to be universally quantified conditionals and biconditionals—*All A's are B's*/$\forall x \, (Ax \supset Bx)$, *All and only A's are B's*/$\forall x \, (Ax \leftrightarrow Bx)$?

(iv) Are theoretical identity sentences involving rigid natural kind predicates guaranteed to be necessary if true?

(v) Is there any semantic property of natural kind predicates, other than rigidity, that guarantees theoretical identity sentences involving predicates possessing that property are necessary if true? If so, are the necessary truths expressible by such sentences knowable a priori, or only a posteriori?

(vi) What are the most important ways in which natural kind predicates are semantically similar to proper names?

Extending Kripke's semantic model from names to natural kind predicates in order to answer these questions is the second important piece of unfinished semantic business left to us by *Naming and Necessity*. This will be my task in chapters 9–11.

2

Rigid Designation and Its Lessons

for the Semantic Contents

of Proper Names

This chapter is concerned with Kripke's doctrine that proper names are rigid designators, and the challenges it poses to analyses that treat the meanings of names as given by definite descriptions associated with them by speakers. We will begin with a brief review of Kripke's arguments against descriptivism about names, and with the isolation of one of these arguments—the so-called modal argument—as depending crucially on the claim that names are rigid. After an explanation of this claim and its role in the modal argument, the bulk of the chapter will be devoted to examining and refuting the leading attempts to circumvent the argument and reinstate descriptivism. The main lesson to be drawn from these failures is that the considerations underlying Kripke's original modal argument can be strengthened and extended so as to constitute a decisive objection to all standard forms of descriptivism about the meaning of names, no matter how sophisticated or convoluted. This objection invites the conclusion that no proper names have the semantic contents of definite descriptions. However, it does not quite establish it. Although the conclusion does seem to hold for the great majority of proper names that have attracted the attention of philosophers, at the end of the chapter we will discuss a distinctive class of proper names that may well constitute a special and highly restricted exception to it.

Three Arguments Against the Descriptivist Picture

In *Naming and Necessity*, Saul Kripke gives three types of argument against semantic theories that analyze the meaning of proper names,

and the manner in which their reference is determined, in terms of the meaning, or denotation, of descriptions associated with those names by speakers. The first type consists of **semantic arguments** designed to show that, typically, the referent of a proper name n, as used by a speaker s, is not linguistically determined to be the denotation of any description, or set of descriptions, associated with n by s. The second type consists of **epistemic arguments** designed to show that what is known or believed by someone who knows or believes that which is expressed by a sentence s containing a proper name n is different from what is known or believed by someone who knows or believes that which is expressed by a sentence which results from substituting a description for n in s. The third type consists of **modal arguments.** These are intended to show that sentences containing names typically have different truth conditions than corresponding sentences containing descriptions, in the sense that sentences of these two types are typically true in different possible states of affairs.

One of Kripke's semantic arguments is based on the observation that in some cases a speaker's use of a name n may uniquely refer to an object o, even though the speaker has no uniquely denoting description at all associated with n. Names of famous people of whose accomplishments most speakers are only dimly aware provide examples of this type. For example, many people have heard the name *Cicero* and know that it refers to a famous Roman, but know little else about him. Nevertheless, such speakers can use the name to refer to a specific man, even though they are not able to provide any description that picks him out uniquely.

The reason they are able to do this, as Kripke points out, is that the linguistic mechanism determining the reference of a speaker's use of a name is typically the historical chain of transmission in which the speaker stands. The standard case goes roughly as follows: A name is introduced and, once introduced, is passed from one speaker to another. Each time it is passed to a new speaker, the person acquiring the name intends to use it to refer to whomever or whatever that person's sources use it to refer to. Often when this happens, the person acquiring the name picks up substantial information about its referent in the process. However, this is not always so, and in some cases considerable misinformation may be passed along. Because of this, speakers' answers to the question

Q.   To whom or what are you using the name n to refer?

are not always reliable. As Kripke has shown, there are cases in which speakers use a name n to refer to an object o even though the descriptions elicited by Q (a) do not pick out any object uniquely,[1] or (b) pick out some unique object other than o.[2] He takes cases like these to refute descriptive theories that claim the referent of an arbitrary proper name n, as used by a speaker s, is linguistically determined to be the unique object (if there is one) satisfying the descriptions that s takes to be definitive of s's use of n.

Some descriptivists have objected that Kripke's conclusion is premature. Although they agree that his semantic arguments show that in most cases the referent of a name for a speaker is not fixed by the descriptions the speaker would most readily give in answer to Q, they insist that there may be other descriptions that fill the bill. Consider, for example, Kripke's own theory about the historical chain of transmission by which reference is normally determined. Surely that theory could be put in the form of a description. But then, if the theory is right, that description fixes reference. And so, it might be claimed, descriptivism is vindicated after all.

However, things are not so simple. First, the historical-transmission account of reference sketched by Kripke leaves many questions unanswered, and falls short of being a complete and explicit theory. For example, we know that sometimes a name is introduced with a certain referent, is passed on to others, and at some stage in the historical chain of transmission loses its initial referent and acquires a new one, without anyone in the chain intending to change the reference of the term.[3] There is nothing in Kripke's discussion that explains precisely how this happens, or that specifies the conditions that have to be met in order for it to occur. This does not, of course, falsify his guiding idea. However, as he was the first to admit, it does show that his idea does not amount to a fully explicit theory that accounts for all instances of a speaker's use of a name referring to an object.[4] Consequently, no description extracted from it constitutes the linguistic mechanism by which the referents of names are definitively determined.

Second, even if one had a complete, explicit theory from which one could extract a definitive reference-fixing description, in order to vindicate descriptivism one would still have to show that ordinary speakers somehow possess this description, and use it to establish the

references of names. This problem is by no means trivial. Surely ordinary speakers cannot produce on demand a fully accurate and explicit description that covers all cases. Moreover, even in the unlikely event that they could somehow be shown to implicitly grasp such a description, this would not be enough. Speakers often have many descriptions associated with a name they use. What, if anything, makes one of those descriptions privileged, in the sense not only of managing to apply to what the name really refers to, but also of playing the central role in determining the reference of the name in the first place? Unless the descriptivist can answer this question, there is no vindication of descriptivism.

Finally, the question of how the reference of names is fixed is less important philosophically than related questions about the meaning or semantic content of names, and of sentences containing them. Kripke approaches these questions by investigating the epistemic and modal properties of sentences containing names, and using these properties to argue that the most straightforward versions of descriptivism about the semantic contents of names can't be correct. Unless these arguments can be answered, there will be little to recommend descriptivism.

This brings us to Kripke's **epistemic** arguments against descriptive theories of names. These arguments are designed to show that the epistemic status of (the propositions semantically expressed by) sentences containing names typically is different from the epistemic status of (the propositions semantically expressed by) corresponding sentences containing descriptions. For example, if D is the description associated with a name n by speakers, then the proposition semantically expressed by the sentence *if n exists, then n is D* (or *if there is/was such a thing as n, then n is/was D*) is typically not knowable a priori even though the proposition expressed by *if D exists, then D is D* (or *if there is/ was such a thing as D, then D is/was D*) is knowable a priori. This supports the conclusion that D does not, in fact, have the same meaning (semantic content) as n.

One example of this type is provided by the name *Christopher Columbus* and the description *the first European to discover America*. Although this description represents the most important thing that most people think about Columbus, the claim that if there was such a person as Columbus, then Columbus was the first European to discover America clearly rests on empirical evidence, and thus is the sort of proposition that could, in principle, be shown to be false by further