

## PART SEVEN

SAUL KRIPKE ON NAMING AND NECESSITY

# CHAPTER 14

## NAMES, ESSENCE, AND POSSIBILITY

### CHAPTER OUTLINE

1. *The Significance of Naming and Necessity*

2. *Why Descriptions Do Not Give the Meanings of Names*

The modal argument that the meanings of names are not given by descriptions speakers associate with them

3. *Rigid Designation*

Rigid Designation and the Modal Argument

Definition of rigid designation; Kripke's argument that names are rigid, whereas descriptions typically associated with them are not; the attempt to avoid this argument by appealing to rigidified descriptions; extension of the modal argument to rebut this proposal

A confusion to be avoided

How it is that the claim that *n* designates *o* may fail to be necessary, even though *n* designates *o* with respect to all possible world-states

Rigid designation and essentialism

The use of rigid designation to rebut Quine's objection that it makes no sense to say of *o*, independently of how it is described, that it has a property essentially

Rigid designation, possible worlds, and "transworld identification"

Possible worlds are not alternate universes requiring criteria for identifying counterparts of actually existing objects; they are possible states the universe could have been in; the sense in which rigid designators are used to "stipulate" these states

4. *Why Descriptions Normally Do Not Fix the Referents of Names*

Arguments against the weak, reference-fixing, version of the description theory

Kripke's historical chain conception of reference determination

Do historical chains provide reference-fixing descriptions?

Why they do not; resolution of an unclarity about reference-fixing

### **The Significance of *Naming and Necessity***

In this chapter, we begin our discussion of Saul Kripke's book *Naming and Necessity*, which was originally presented as three long public lectures at Princeton University in January of 1970, when Kripke was 29

years old.<sup>1</sup> A tape recording was made of the lectures, and two professors in the Princeton philosophy department at the time, Gilbert Harman and Thomas Nagel, produced a transcript from the tapes. Kripke added footnotes, and later wrote a preface for the book version. The impact of the lectures was profound and immediate, and over the years their influence has grown. In the philosophy of language, *Naming and Necessity* is among the most important works ever, ranking with the classical work of Frege in the late nineteenth century, and of Russell and Tarski in the first half of the twentieth century. Beyond the philosophy of language, it fundamentally changed the way in which much philosophy is done. The most important aspects of the work are (i) a set of theses about the meaning and reference of proper names, (ii) a corresponding set of theses about the meaning and reference of natural kind terms such as *heat*, *light*, *gold*, *water*, and *tiger*, (iii) a compelling defense of the metaphysical concepts of necessity and possibility, (iv) a sharp distinction between the metaphysical notion of necessity and the epistemological notion of apriority, (v) forceful arguments that there are necessary truths that are knowable only a posteriori, and apriori truths that are contingent, and hence not necessary, and (vi) a persuasive defense of the intelligibility of essentialism—i.e., the claim that it makes sense to characterize objects as having some of their properties essentially, and others accidentally. In addition to these explicit aspects of the work, the discussion in *Naming and Necessity* had far-reaching implications for what has come to be known as *externalism* about meaning and belief—roughly, the view that the meanings of one's words, as well as the contents of one's beliefs, are partly constituted by facts entirely outside oneself. Finally, *Naming and Necessity* played a large role in the implicit, but widespread, rejection of the view—so popular among ordinary language philosophers—that philosophy is nothing more than the analysis of language.

### Why Descriptions Do Not Give the Meanings of Names

We will begin our investigation of Kripke's work by looking at his discussion of the description theory of proper names, of which he distinguishes two different versions. According to the first, proper names have the same meanings as descriptions that speakers associate with

<sup>1</sup> Saul Kripke, *Naming and Necessity*, Cambridge, MA: Harvard, 1980, originally published in D. Davidson and G. Harman, eds., *Semantics of Natural Languages* (Dordrecht: Reidel, 1972); citations will be to the 1980 text.

them. According to the second, although names may not be synonymous with descriptions, the referent of a proper name *n*, as used by a speaker at a given time, is determined, as a matter of linguistic rule, to be the one and only object that satisfies the descriptions associated with *n* by the speaker at that time. Since the meaning of a term is assumed to determine its reference, the first version of the description theory is understood as entailing the second. However, the converse does not hold; it may turn out that the referent of a name is, as a matter of semantic rule, determined by a description, even though the name is not synonymous with the description. These two versions of the description theory are expressed by theses 1 and 2.<sup>2</sup>

#### TWO THESES ABOUT NAMES

Thesis 1: The meaning of a name *n* (for a speaker at a time) is given by a description, a conjunction of descriptions, or a cluster of descriptions *D* that the speaker associates with *n* at the time. If *D* gives the meaning of *n*, then substitution of one for another preserves both meaning and proposition expressed. Thus, if *S'* results from *S* by substitution of *D* for one or more occurrences of *n* in *S*, then *S* and *S'* mean the same thing, and express the same proposition.

Thesis 2: The referent of a name *n* (for a speaker at a time) is semantically fixed (determined) by a description, a conjunction of descriptions, or a cluster of descriptions *D* that the speaker associates with *n* at the time. If *D* fixes the referent of *n*, then:

- (i) the speaker believes that *D* applies to a unique individual;
- (ii) if *D* does apply to a unique individual *o*, then *o* is the referent of *n*;
- (iii) if *D* does not apply to a unique individual, then *n* has no referent.
- (iv) the speaker knows (or is capable of knowing) apriori that *if n exists, then n is D* expresses a truth.

In lecture 1, Kripke gives an argument against thesis 1, which has come to be known as *the modal argument*. We illustrate the argument

<sup>2</sup> These are reconstructions of the two theses as Kripke understands them. For Kripke's own formulations, see pp. 71–80 of *Naming and Necessity*.

with an example. Let  $n$  be the name *Aristotle*. Let (1–8) be candidates for the description  $D$  that gives the meaning of the name.

1. the founder of formal logic
2. the greatest student of Plato
3. the teacher of Alexander
4. the famous Greek philosopher named ‘Aristotle’
5. the last great philosopher of antiquity
6. the conjunction of 1–5
7. the conjunction of all descriptions the speaker associates with  $n$
8. a cluster of descriptions including 1–5 that the speaker associates with  $n$

For our purposes, the claim that the meaning of a name is given by a cluster of descriptions  $D_1 \dots D_n$  will be regarded as equivalent to the claim that the meaning of the name is given by the description *the thing of which most (or a sufficient number) of the following are true: it is  $D_1$ , it is  $D_2$ , . . . , it is  $D_n$* . We may now test the claim that the meaning of *Aristotle* is given by one or more of the descriptions (1–8), by applying the following, modal test.

#### THE MODAL TEST

If  $D$  gives the meaning of  $n$ , then the proposition (truth/statement/claim) expressed by the sentence

*If  $n$  existed, then  $n$  was  $D$ .*

is a necessary truth.

The rationale behind this test is that if  $D$  has the same meaning as  $n$ , then substitution of one for the other in a sentence won’t change the proposition expressed (or the statement made). But that means that the sentence *If  $n$  existed, then  $n$  was  $D$*  expresses the same proposition (says the same thing) as the sentence *If  $D$  existed, then  $D$  was  $D$* . Since the latter sentence expresses a necessary truth, so does the former sentence. In using the terminology *necessary truth* we mean the following: A proposition is a necessary truth iff (i) it is true given the way the world actually is and (ii) it would have been true had the world been in any other possible state it could have been in.

If thesis 1 is correct, then there must be some description *D* which speakers associate with the name ‘Aristotle’ such that the proposition expressed by *If Aristotle existed, then Aristotle was D* is a necessary truth. In fact, since *D* gives the meaning of ‘Aristotle’, the proposition expressed by that sentence should be both necessary and knowable a priori. But, Kripke argues, there is no such description *D*. For example, consider the description *the founder of formal logic* as a possible candidate. To apply the modal test, we ask whether the statement *If Aristotle existed, then Aristotle was the founder of formal logic* is a necessary truth. To say that it is a necessary truth is to say that there is no possible way that the world could have been which would make the antecedent, *Aristotle existed*, true, but the consequent, *Aristotle was the founder of formal logic*, false. But that doesn’t seem right. On the contrary, it seems that the world could have been in a state in which Aristotle existed, but didn’t do any logic at all. Since Aristotle could have existed without being the founder of formal logic, the conditional statement *If Aristotle existed, then Aristotle was the founder of formal logic* is not a necessary truth. Thus, the name *Aristotle* does not mean the same thing as the description *the founder of formal logic*.

This result is not an isolated one. The same argument could be given for the other candidate descriptions (1–8), or for other descriptions with which one would most naturally think of replacing the name *Aristotle*. The reason for this is that nearly all descriptions people nowadays associate with the name have to do with Aristotle’s famous accomplishments, especially those involving his philosophy. However, as Kripke rightly points out, none of those accomplishments were necessary conditions for Aristotle to exist.<sup>3</sup> Aristotle could have existed even if he hadn’t gone into philosophy, or done anything important. Thus, Kripke is prepared to run the argument we have just given using any description or cluster of descriptions based on Aristotle’s famous achievements, or well-known characteristics. But since these provide the main descriptive content that most of us associate with the name, he concludes that the description theory of meaning is incorrect as a theory of how most of us use the name.

The force of this argument against the description theory of the meaning of proper names can be brought out by asking what conditions a description would have to satisfy in order to block the argument. At a minimum, the description *D* would have to be such that the proposition expressed by *If n existed, then n was D* was both necessary and a priori.

<sup>3</sup> Even being named ‘Aristotle’ was not a necessary condition for Aristotle to exist.

The problem is that for nearly all names one can think of, one can't come up with such descriptions. If this is right, then thesis 1 is false.

The argument we have just given against thesis 1 is **not** directed against thesis 2, which expresses the view that descriptions semantically fix the referents of names. The reason for this is that thesis 2 does **not** claim that names and descriptions are synonymous. It was only the claim that names have the same meanings as certain descriptions that allowed us to view thesis 1 as committed to the claim that when we substitute the names and descriptions for one another in a sentence, we do not change the proposition expressed, and so do not change the modal or epistemological status of the sentence. Suppose, however, that we had a theory that told us not that D gives the meaning of n, but only that D semantically fixes the referent of n. Such a fix-the-referent theory would **not** claim that n is synonymous with D, and so would **not** be committed to the claim that substitution of one for another must preserve the proposition expressed. But if substitution could change the proposition expressed, then we have no reason to think that it couldn't also change the modal or epistemic status of the sentence. Thus, although the modal argument may be seen as refuting the view that names are synonymous with the descriptions that speakers associate with them, it does not, by itself, refute the view that these descriptions semantically determine the referents of names. So, if one wants to criticize the fix-the-referent version of the description theory, one must come up with additional arguments. In lectures 2 and 3 Kripke does this. But before we look at those arguments, we need to examine some of the key concepts he employs.

## Rigid Designation

### *Rigid Designation and the Modal Argument*

The first such concept is that of a rigid designator.<sup>4</sup>

#### RIGID DESIGNATION

A singular term t is a rigid designator of an object o iff t designates o with respect to all possible states of the world (in

<sup>4</sup> This is a simplified definition that abstracts away from various complications. For example, we here leave indexicals and variables out of consideration, and so need not relativize reference to contexts and assignments of values to variables. For a definition taking these complications into

which *o* exists); and, moreover, *t* never designates anything other than *o* (with respect to any possible-world-state).

If a singular term *t* is a rigid designator of an object *o*, then sentences containing *t* are true when taken as descriptions of alternative possible states  $w_1, w_2, w_3$  of the world iff one and the same object, *o*, has the relevant properties in those alternative states. For example, if *t* is a rigid designator of an object *o*, *F* expresses the property  $\phi$ , and *p* is the proposition expressed by *Ft*, then (i) *p* is true relative to the actual state of the world iff as things actually are in the world, *o* has the property  $\phi$ , and (ii) *p* is true with respect to any other possible state *w* of the world iff relative to *w*, *o* has property  $\phi$  (i.e., if the world were in state *w*, then *o* would have property  $\phi$ ). If *t* is a non-rigid designator of *o*, then although (i) remains as above, (ii) does not. If *t* is non-rigid, then there are sentences *Ft*, properties  $\phi$ , propositions *p*, and possible states of the world *w* and  $w^*$ , such that *p* is expressed by *Ft*, and either *p* is true with respect to *w* even though *o* does not have  $\phi$  with respect to *w*, or *p* is false with respect to  $w^*$  even though *o* has  $\phi$  with respect to  $w^*$  (or both).

For example, consider the pair (9) and (10).

9. The winner of the 1996 U.S. presidential election was a Democrat.

10. Bill Clinton was a Democrat.

The description *the winner of the 1996 U.S. presidential election* and the name *Bill Clinton* designate the same individual *o*. Since *o* was a Democrat, both (9) and (10) are true, with respect to the world as it actually is (was). This is not so with respect to a possible state *w* of the world in which Clinton runs as a Democrat but the Republican Bob Dole wins the 1996 election. The individual whose party affiliation determines the truth value of (10) relative to the world-state *w* is Bill Clinton, the same individual who is relevant to determining the truth value of (10) in the actual world-state. Hence, (10) is true with respect to *w*. However, the individual whose party affiliation determines the truth value of (9) relative to *w* is not Bill Clinton, but rather is Bob Dole. Hence (9) is false with respect to *w*. This example shows that the

---

account, see my *Reference and Description: The Two-Dimensionalist Attempt to Revive Descriptivism* (in preparation). For present purposes, we take singular terms to include names and singular definite descriptions like *the so and so*.



description *the winner of the 1996 U.S. presidential election* is non-rigid, whereas the name *Bill Clinton* is rigid.

How does it happen that for a rigid designator the same object *o* is relevant with respect to all possible states of the world; whereas for a non-rigid designator the relevant objects differ from world-state to world-state? The answer may be expressed as follows: If *t* is rigid, then whatever object is actually designated by *t* is designated by *t* with respect to all possible states of the world in which that object exists, and nothing other than that object is designated by *t* with respect to any world-state; but if *t* is non-rigid, then either what is actually designated by *t* fails to be designated by *t* with respect to other possible world-states (in which that object exists), or something other than that object is designated by *t* with respect to some world-state.<sup>5</sup> This suggests a linguistic test for determining whether an arbitrary singular term in English is a rigid designator.

#### A LINGUISTIC TEST

*t* is a rigid designator iff the sentence *The individual that is (was) actually t could not have existed without being t, and nothing other than the individual that is (was) actually t could have been t* expresses a truth.

Alternatively put: A singular term *t* of English is a rigid designator iff the relevant sentences of the form (11) and (12) are false. A singular term is non-rigid iff either (11) or (12) is true.

11. The individual who is (was) actually *t* could have existed without being *t*.
12. It could have been the case that someone other than the individual who is (was) actually *t* was *t*.

Kripke maintains that if we apply this test we will find that proper names are rigid designators, whereas most ordinary descriptions are not. This is not to say that he thinks that no descriptions are rigid; for example, he would recognize *the positive square root of 25* and *the individual who is identical with Saul Kripke* to be rigid. However, he believes that where most ordinary names are concerned, descriptions like these are not good candidates for being ones used by speakers to give

<sup>5</sup> This idea is a rough approximation which ignores certain abstruse complications that arise in special cases when indexicals are considered and the notion of rigid designation is relativized to context. (The same is true of the linguistic test that follows.)

meanings or establish reference.<sup>6</sup> So, in the case of the vast majority of ordinary names, he would maintain that the descriptions associated with them are non-rigid. In particular, the descriptions (1–8) associated with the name *Aristotle* in our example are non-rigid.

With this in mind, we can restate Kripke's modal argument. In essence, the argument is the following:

P1. Names are rigid designators.

P2. Standardly, the descriptions associated by speakers with names are not rigid designators.

---

C. So, names are standardly not synonymous with descriptions associated with them by speakers.

This argument was immediately recognized to be a powerful challenge to descriptivism about the meanings of proper names, and it continues to be accepted by a great many philosophers to this day.

In recent years however, a response to the argument has gained a certain amount of currency among some theorists who continue to be attracted to descriptivism. The response is based on the observation that for any non-rigid description *the F* we can form a rigid description *the actual F* that designates, when we are talking about any possible world-state, the individual that is F in the actual world-state.<sup>7</sup> Consider, for example, sentence (13).

13. It could have been the case that the actual winner of the 1996 presidential election didn't win the 1996 presidential election.

Intuitively, what this sentence says is true. This means that there must be a possible world-state *w* such that sentence (14), as used by us, is true when taken as a description of *w*.

14. The actual winner of the 1996 presidential election didn't win the 1996 presidential election.

This, in turn, means that the individual who counts as the referent of our use of the term *the actual winner of the 1996 presidential election*,

<sup>6</sup> We here leave aside special cases like the numeral '2', which conceivably might be defined *the successor of 1*.

<sup>7</sup> This statement is somewhat of an exaggeration, and ignores the complications mentioned in footnotes 4 and 5. Those interested in a fuller story should see the discussion of the use of the actuality operator to rigidify descriptions in the work cited in footnote 4.

when used to say something about *w*, is the individual who won the election, not with respect to *w*, but with respect to the world as it actually is. Sentence (14) is true with respect to *w* because that individual—the person who was the winner with respect to the actual world-state—didn't win, with respect to *w*. The lesson here is that the result of adding the actuality operator to a description *the F* results in a new description, *the actual F*, which rigidly designates the object that the former description designates in the actual world-state (if such an object is uniquely designated by *the F* relative to the actual state of the world). This idea has been used by some post-Kripkean descriptivists to suggest that names are synonymous, not with ordinary descriptions, but with descriptions rigidified using the actuality operator. This claim in effect attacks premise P2 above, and is **not** refuted by the modal argument that Kripke gives.

Nevertheless, the proposal fails for other reasons. First, if the proposal were correct, then the proposition expressed by *If *n* existed, then *n* was *D** would be the same as the proposition expressed by *If the actual *D* existed, then the actual *D* was *D**. This latter proposition is something knowable apriori, independent of any empirical evidence. However, when *n* is an ordinary proper name, the proposition expressed by *If *n* existed then *n* was *D** typically is not knowable apriori. Thus, the two propositions are different, and the proposal for saving thesis 1 founders. (This point is implicit in *Naming and Necessity*. We will return to it later, when we discuss the material in Kripke's lecture 2.)

There is also a second problem with the proposal that Kripke did not address, but which is discussed at length in chapter 2 of my *Beyond Rigidity*.<sup>8</sup> I will mention only the gist of it. The proposition that the actual *F* is *G* is a proposition that says of the actual state of the world that the individual who is uniquely *F* with respect to it is also *G*.<sup>9</sup> Consequently, it is possible to believe that the actual *F* is *G* only if one is in some kind of epistemic contact with the actual state of the world, and, by virtue of this, believes a certain proposition about it. Let us suppose, for the sake of argument, (i) that all of us, living in the world as it actually is, do have such contact with this state of the world, and (ii) that because of this, when we believe that the *F* is *G* we automatically qualify as **also** believing, of the actual state of the world, that the individual who is uniquely *F* with respect to it is *G* as well. Even if we grant this, however, we must also recognize that things change when

<sup>8</sup> Scott Soames, *Beyond Rigidity* (New York: Oxford University Press, 2002).

<sup>9</sup> In this paragraph, I use 'F' and 'G' as schematic letters. 'D' is used as a metalinguistic variable.

we consider, not what agents actually believe, but what it is possible for agents to believe. Certainly, it is possible to believe that Aristotle was a philosopher without being in any epistemic contact with the actual world-state, and hence without believing anything about it. For example, if various facts irrelevant to Aristotle had been different, I might still have believed that Aristotle was a philosopher. To say this is just to say that there is some possible world-state  $w$ , different from the actual world-state, such that with respect to  $w$  I believe that Aristotle was a philosopher, even though with respect to that world-state I may not believe anything about the actual world-state. I may not believe anything about the actual world-state, since the actual world-state is a total or maximal property which represents the way things actually are. Had the world been in state  $w$ , I may have been familiar with the way things are with respect to  $w$ , and so have been epistemically acquainted with  $w$ , but I need not have been familiar with other maximally possible states of the world, and so I need not have been familiar with the actual world-state. If this is right, it shows that for any description *the actual*  $D$ , it is possible to believe the proposition expressed by *Aristotle was a philosopher* without believing the proposition expressed by *The actual  $D$  was a philosopher*, and hence that the two propositions are different.<sup>10</sup> Thus, the proposal that names are synonymous with descriptions rigidified using the actuality operator is false.

### *A Confusion to Be Avoided*

Before going further, we pause to dispel a confusion that is all too easy to fall into, but very important to avoid. The confusion arises from a puzzle generated by the following two claims.

- (i) The name ‘Aristotle’ is a rigid designator. Thus, for all possible states of the world  $w$ , it refers to the same individual—the man Aristotle—with respect to  $w$ .
- (ii) It is not a necessary truth that Aristotle was named ‘Aristotle’; it could have been the case that the name ‘Aristotle’ did not

<sup>10</sup> This step in the argument tacitly assumes, as did the previous argument against the proposal, that  $x$  believes that  $S$  reports a relation between the believer and the proposition expressed by  $S$ . Some proponents of analyzing names as descriptions rigidified using the actuality operator dispute this assumption. For rejoinders to the most systematic attempt to develop such a point of view, see my “Saul Kripke, the Necessary Aposteriori, and the Two-Dimensionalist Heresy,” in M. Garcia-Carpintero and J. Maciá, eds. *The Two-Dimensional Framework: Foundations and Applications* (Oxford: Oxford University Press, forthcoming), and my *Reference and Description: The Two-Dimensionalist Attempt to Revive Descriptivism*, in preparation.

refer to Aristotle. Thus, there must be some world-state  $w$  such that the claim that ‘Aristotle’ does not refer to Aristotle is true with respect to  $w$ .

Both claims are true, and both would be endorsed by Kripke. However, this might seem puzzling, since (i) and (ii) may appear to be inconsistent. What makes (i) and (ii) seem inconsistent is the tendency to tacitly accept (iii) as something so obvious as to go without saying.

- (iii) The three-place relation \_\_\_\_\_ *refers to* \_\_\_\_\_ *with respect to* \_\_\_\_\_ (tacitly invoked in (i)) holds between the name ‘Aristotle’, the man Aristotle, and a world-state  $w$  iff it is true with respect to  $w$  that the two-place relation \_\_\_\_\_ *refers to* \_\_\_\_\_ (invoked in (ii)) holds between ‘Aristotle’ and Aristotle—i.e., iff the claim that ‘Aristotle’ refers to Aristotle is true when taken as a description of  $w$ .

Although principle (iii) might seem undeniable at first glance, in fact it is false. Throughout this chapter we have been following Kripke in taking the three-place relation \_\_\_\_\_ *refers to* \_\_\_\_\_ *with respect to* \_\_\_\_\_ to hold between a name  $n$ , object  $o$ , and world-state  $w$  iff  **$n$ , as used by us here and now in the world as it actually is, refers to the object  $o$ , when our words are taken as descriptions of  $w$** . Because of this,  $n$  may refer to  $o$  with respect to  $w$  even if (a) the name  $n$  doesn’t exist with respect to  $w$ , or (b) with respect to  $w$ , the name  $n$  is not used by speakers to refer to anything, or (c) with respect to  $w$ , the name  $n$  is used by speakers to refer to something other than  $o$ . What, if anything, speakers would have used the name  $n$  to refer to had the world been in state  $w$  is irrelevant to whether  $n$  refers to  $o$  with respect to  $w$ . However, what speakers would have used  $n$  to refer to, had the world been in state  $w$ , is crucial to determining which pairs of names and objects the two-place relation \_\_\_\_\_ *refers to* \_\_\_\_\_ applies to with respect to  $w$ . **It is true with respect to  $w$  that the name  $n$  refers to the object  $o$  iff had the world been in state  $w$ , speakers would have used  $n$  to refer to  $o$** . Thus, what (ii) says is that there are world-states such that had the world been in those states, speakers would not have used ‘Aristotle’ to refer to Aristotle. This is compatible with the claim made by (i)—namely that, here and now in the world as it actually is, we use the name ‘Aristotle’ to refer to the man Aristotle when our words are taken as descriptions of any world-state whatsoever.

*Rigid Designation and Essentialism*

Throughout Kripke's discussion of names in lecture 1 of *Naming and Necessity*, he takes it for granted that the distinction between the essential properties of an object and its contingent properties is a legitimate one. An essential property of an object is a property the object could not lack in any circumstance in which it existed at all. A contingent, or accidental, property is one that the object has, but could have existed without. Examples of contingent properties of mine are the property of living in Princeton, the property of being a father, and the property of being a philosopher. Uncontroversial examples of essential properties of mine are rarer, but the following seem to be good candidates: the property of being a human being, the property of having a brain, the property of having a body made up of molecules, the property of being mortal, the property of not being identical with Saul Kripke.

There is a close connection between the notion of a rigid designator and the claim that an object has a property essentially. This connection is expressed by (i).<sup>11</sup>

- (i) If *n* is a rigid designator of *o*, and *F* is a predicate expressing the property *P*, then the claim that *P* is an essential property of *o* is equivalent to the claim *it is necessary that if n exists, then n is F*.

The equivalence mentioned in (i) is relevant to a puzzle that had been used by Quine for more than two decades prior to *Naming and Necessity* to cast doubt on the intelligibility of essentialism.<sup>12</sup> Quine had contended first, that the notion of an essential property of an object is

<sup>11</sup> Some explanation of my terminology may be in order. Predicates **express** properties while being **true of** objects with respect to world-states. Note, what property a predicate expresses is not relativized to different world-states. If *F* expresses *P*, then for any world-state *w*, *F* is true of an object *o* with respect to *w* iff *o* has *P* with respect to *w*.

<sup>12</sup> See Quine, "Notes on Existence and Necessity," *Journal of Philosophy* 40 (1943): 113–27; "The Problem of Interpreting Modal Logic," *Journal of Symbolic Logic* 12 (1947): 43–48; "Reference and Modality," in *From a Logical Point of View* (Cambridge, MA: Harvard University Press, 1953, 1961, 1980); and "Three Grades of Modal Involvement," originally published in 1953, reprinted in Quine, *The Ways of Paradox* (New York: Random House, 1966). For explication and critique of Quine, each excellent in its way, see David Kaplan, "Opacity," in E. Hahn and P. A. Schilpp, eds., *The Philosophy of W. V. Quine* (La Salle, IL: Open Court, 1986), and John Burgess, "Quinus ab omni naevo vindicatus," in Ali A. Kazmi, ed., *Meaning and Reference* (Calgary: University of Calgary Press, 1998). Burgess makes the point that if one (wrongly) regards necessity simply as analyticity, and takes analyticity to be a property of sentences—things which both Quine and those he was criticizing at the time did—then his skepticism about "analytically essential properties of an object" makes much more sense (though it no longer has anything to do with genuine metaphysical

defined in terms of the notion of necessity, and second, that whatever grasp we have of necessity is expressed by our use of the predicate *is a necessary truth*, which applies to sentences, or by our use of the operator *it is necessary that . . .*, which attaches to sentences. Thus, he thought, if we are to make sense of the idea that *is essentially F* applies to a certain object *o*, we have no choice but to view this claim as based on the assumption that the sentence *it is necessary that if t exists, then t is F* is true, for some designated choice of a term *t* that refers to, or describes, *o*.

However, Quine also observed that for any object *o*, there will be some terms *t* that refer to *o* which make the sentence *It is necessary that if t exists, then t is F* false, even if there are other terms *t* that refer to *o* which make it true. Thus, he thought, relative to one way of describing *o* it may turn out that the property expressed by *F* is an essential property of *o*, whereas relative to a different way of describing *o* it will turn out that the property expressed by *F* is not an essential property of *o*. But what if we consider *o* on its own, independent of any description? Is the property expressed by *F* one of *o*'s essential properties or not? It would seem that there is no saying.

Here is a representative example of Quine's presentation of this allegedly puzzling problem.

Perhaps I can evoke the appropriate sense of bewilderment as follows. Mathematicians may conceivably be said to be necessarily rational and not necessarily two-legged; and cyclists necessarily two-legged and not necessarily rational. But what of an individual who counts among his eccentricities both mathematics and cycling? Is this concrete individual necessarily rational and contingently two-legged or vice versa? Just insofar as we are talking referentially of the object, with no special bias toward a background grouping of mathematicians as against cyclists or vice versa, there is no semblance of sense in rating some of his attributes as necessary and others as contingent. Some of his attributes count as important and others as unimportant, yes; some as enduring and others as fleeting; but none as necessary or contingent.<sup>13</sup>

---

necessity, or with metaphysically essential properties). Kaplan argues (i) that Quine's position on the matter was tied to a larger mistaken opposition to "quantifying in" to all nonextensional constructions, (ii) that anyone who accepts logical truth can accept benign versions of essentialism, and (iii) that though the intelligibility of more robust essentialist claims should not be in doubt, the truth, or the falsity, of such claims does raise substantive metaphysical questions.

<sup>13</sup> Quine, *Word and Object*, p. 199.

Let *i* be some individual who is both a brilliant mathematician and a champion cyclist, and suppose that *the world's greatest mathematician* and *the world's greatest cyclist* both designate *i*. Then, since (15a) is, arguably, true, while (15b) is not,

- 15a. It is necessary that: if the world's greatest mathematician exists (i.e., if there is such an individual as the world's greatest mathematician), then the world's greatest mathematician is rational.
- b. It is necessary that if the world's greatest mathematician exists (i.e., if there is such an individual as the world's greatest mathematician), then the world's greatest mathematician is two-legged.

it follows, on Quine's view, that relative to the choice of describing *i* as *the world's greatest mathematician*, being rational is one of *i*'s essential properties, but being two-legged is not. However, if we choose to describe *i* as *the world's greatest cyclist*, we get the opposite result. Since (16a) is, arguably, true, while (16b) is not,

- 16a. It is necessary that: if the world's greatest cyclist exists (i.e., if there is such an individual as the world's greatest cyclist), then the world's greatest cyclist is two-legged.
- b. It is necessary that: if the world's greatest cyclist exists (i.e., if there is such an individual as the world's greatest cyclist), then the world's greatest cyclist is rational.

it follows, on Quine's view, that relative to the choice of describing *i* as *the world's greatest cyclist*, being two-legged is one of *i*'s essential properties, but being rational is not. Accordingly, Quine thinks, it makes no sense to ask of *i* himself, independent of any way of describing him, which of his properties are essential and which are not.

In general, Quine assumed that there was no principled, non-arbitrary way of selecting, for an arbitrary object *o* and property *P*, what sort of term *t* should be used to underwrite claims to the effect that *o* did, or did not, have *P* essentially. Thus, his doctrine was that it makes no sense to ask whether an object has a property essentially, independently of how it is described. Rather, objects have or lack properties essentially only **relative** to ways of describing them.

The relevance of all this to Kripke is that if, as he maintains, there is a genuine distinction between rigid and non-rigid designators, then



rigid designators provide a principled connection between claims about the essential properties of objects and claims about which sentences do, and which sentences do not, express necessary truths. It is only sentences of the sort mentioned in (i) that contain a **rigid** designator of the object *o* that are relevant to the question of whether *o* has the property *P* essentially or not. When we consider whether an object has a property essentially, we use a rigid designator to talk about one and the same object with respect to all possible world-states. Because the designator is rigid, the question of whether the object has the property with respect to all the world-states is equivalent to the question of whether the sentence or formula attributing the property to the referent of the rigid designator is true relative to all those world-states. The truth values of other sentences containing non-rigid designators of the object are simply irrelevant.<sup>14</sup>

In this way, Kripke rebuts Quine's objection to the intelligibility of essentialism. The dialectical situation is this: We begin with an intuitive distinction. Although I am correctly described as the Princeton philosopher who was raised in Seattle, being a philosopher, working at Princeton, and being raised in Seattle are contingent properties of mine—I could have existed even if I hadn't been raised in Seattle, gone into philosophy, or worked at Princeton. By contrast, being a sentient being and not being identical with Saul Kripke seem to be essential properties of mine—there seem to be no possible scenarios in which I exist but am not a sentient being, or in which I am Saul Kripke. Everyone understands these claims. Although there may be disagreements and uncertainties about which properties fall into which categories, we all recognize the intelligibility of claims of this sort prior to receiving any instruction in philosophy. Then Quine comes along with an objection. He gives an argument that is designed to show that we have all been unwittingly talking nonsense. However, his objection relies on a false premise—namely, that there is no non-arbitrary way of selecting, for a given object *o* and property *P* expressed by a predicate *F*, what sort of term *t* designating *o* should be used to construct the statements, *It is necessary that if t exists, then t is F*, upon which the truth or falsity of essentialist claims about *o* depend. Kripke refutes this premise by showing that rigid designators and only rigid designators provide the connection between claims about the necessity of statements, on the one

<sup>14</sup> Note, since *the world's greatest mathematician* and *the world's greatest cyclist* are both non-rigid, the sentences in (15) and (16) are irrelevant to the question of whether or not the individual denoted by them both is essentially rational, or is essentially two-legged.

hand, and the essential properties of objects, on the other. With Quine's objection out of the way, our pre-philosophical conviction that essentialist claims are intelligible stands unchallenged.

At this point, one must be on guard against a certain all too familiar response from unreconstructed Quineans. Of course, they will say, if rigid designation makes sense, then essentialism also makes sense. But does rigid designation make sense? Take the name *Aristotle*, for example. To say that it is a rigid designator is to say that the claim made by our use of a sentence like *Aristotle was a philosopher* is true when evaluated with respect to a possible world-state *w* (e.g., one in which the man *m* we actually call 'Aristotle' never met or studied with Plato) iff *m* was a philosopher (i.e., had the property of being a philosopher) with respect to *w*. But this makes sense only if it makes sense to ask of the individual *m*, independent of any description, whether *m* had a certain property relative to some merely possible world-state. Surely, this is the sort of thing that Quine questioned. Hence, the Quinean skeptic maintains, Kripke's appeal to rigid designation is question-begging, and Quine's powerful objection remains intact.

In my view, and in Kripke's, this signature move of the Quinean skeptic is entirely misguided. We began with an intuitive, pre-philosophical distinction the intelligibility of which is recognized by nearly everyone. Quine offers an objection. He purports to show that there is something incoherent in the way we have all been looking at things. The burden of proof is then on **him** to demonstrate that there is some internal incoherence in our thinking, some conflict between different aspects of our view that we had not noticed. Kripke's rebuttal of Quine's objection shows that he failed to do this. At this point, it is no response to reply that a real Quinean skeptic—one determined, at any cost, not to accept the intelligibility of essentialist claims—would not grant the presuppositions of Kripke's rebuttal. It is not up to Kripke to prove the intelligibility of essentialism from premises that would have to be accepted by even the most determined skeptic, any more than it is the job of the opponent of radical skepticism about the external world to prove the existence of objects other than oneself and one's ideas to the satisfaction of a philosopher determined to take a solipsistic stand. It is sufficient to rebut any reasoned objections that such skeptics may raise in an attempt to persuade us that, by standards that even we must recognize, our ordinary, commonsense views are in error.

In light of this, it is not surprising that Kripke's final response to Quine on this point has a distinctly Moorean flavor. Immediately after

a paragraph summarizing Quine's objection to the intelligibility of the claim that objects have both essential and accidental properties, independent of how they are described, Kripke says the following:

It is even suggested in the literature, that though a notion of necessity may have some sort of intuition behind it (we do think some things could have been otherwise; other things we don't think could have been otherwise), this notion [of a distinction between necessary and contingent properties] is just a doctrine made up by some bad philosopher, who (I guess) didn't realize that there are several ways of referring to the same thing. I don't know if some philosophers have not realized this; but at any rate it is very far from being true that this idea [that a property can meaningfully be held to be essential or accidental to an object independently of its description] is a notion which has no intuitive content, which means nothing to the ordinary man. Suppose that someone said, pointing to Nixon, 'That's the guy who might have lost'. Someone else says 'Oh no, if you describe him as 'Nixon', then he might have lost; but, of course, describing him as the winner, then it is not true that he might have lost'. Now which one is being the philosopher, here, the unintuitive man? It seems to me obviously to be the second. The second man has a philosophical theory. The first man would say, and with great conviction, 'Well, of course, the winner of the election *might have been someone else*. The actual winner, had the course of the campaign been different, might have been the loser, and someone else the winner; or there might have been no election at all. So such terms as "the winner" and "the loser" don't designate the same objects in all possible worlds. On the other hand, the term "Nixon" is just a *name of this man*'. When you ask whether it is necessary or contingent that *Nixon* won the election, you are asking the intuitive question whether in some counterfactual situation, *this man* would in fact have lost the election. If someone thinks that the notion of a necessary or contingent property (forget whether there *are* any nontrivial necessary properties [and consider] just the meaningfulness of the notion) is a philosopher's notion with no intuitive content, he is wrong. Of course, some philosophers think that something's having intuitive content is very inconclusive evidence in favor of it. I think it is very heavy evidence in favor of anything, myself. I really don't know, in a way, what more conclusive evidence one can have about anything, ultimately speaking. But, in

any event, people who think the notion of accidental property un-intuitive have intuition reversed, I think.<sup>15</sup>

At bottom, Kripke's position in the face of the Quinean skeptic about the intelligibility of essentialism is rather like Moore's position in the face of the skeptics he faced.<sup>16</sup> For Kripke, (i) there is a strong initial presumption that both our ordinary counterfactual talk and the distinction between essential and accidental properties that goes along with it are intelligible, and (ii) in order to maintain the distinction, it is enough to rebut skeptical arguments designed to demonstrate that no such distinction could be coherent.

If, as I believe, this is the right way to view the situation, why was Quine's skepticism on this point so influential for so long? In my opinion, three main factors played important roles. First, for many years Kripke's technical apparatus of possible world semantics, including his concept of rigid designation, was either nonexistent, not widely understood, or imperfectly grasped (and sometimes entangled with extraneous and implausible views). Without a useful and readily applicable concept of rigid designation, it was not entirely clear what the response to Quine's skeptical objection should be.<sup>17</sup> Second, as in many discussions of skepticism, issues about who bears the burden of proof became thoroughly confused from very early on, with defenders of Quine (perversely) refusing to grant the initial presumption of intelligibility to the commonplace counterfactual talk appealed to by Kripke and other proponents of the distinction between essential and accidental properties. Third, the by now familiar confusion of necessity with analyticity played a large role in obscuring the central issues at stake.

<sup>15</sup> *Naming and Necessity*, pp. 41–42. The square brackets are Kripke's.

<sup>16</sup> See chapters 1 and 2 of volume 1 for a discussion of Moore's response to skepticism.

<sup>17</sup> There were, to be sure, historical anticipations of Kripke's doctrines about rigid designators, and their potential use in replying to Quine-style objections, including important discussions by Bertrand Russell, Raymond Smullyan, Frederick Fitch, Ruth Barcan Marcus, Paul Ziff, Arthur Prior, Dagfinn Føllesdal, Keith Donnellan, Peter Geach, and others. For historical background, see my "Revisionism about Reference" and "More Revisionism about Reference," along with John Burgess's "Marcus, Kripke, and Names" and "How Not to Write History of Philosophy," all of which can be found in Paul W. Humphreys and James H. Fetzer, eds., *The New Theory of Reference* (Dordrecht, Boston and London: Kluwer, 1998). Although written in response to a nasty controversy involving proper historical credit, these papers contain significant information about Kripke's precursors, and early attempts to reply to Quine. (A different aspect of the controversy is covered by Burgess in his unpublished manuscript, "Geach, Donnellan, Kripke, and Names.") For a useful and extensive survey of a number of relevant topics, see also Stephen Neale, "On a Milestone of Empiricism," in Alex Orenstein and Petr Kotatko, eds., *Knowledge, Language, and Logic* (Dordrecht and London: Kluwer, 2000).

Quine's discussions make clear both that he identified analyticity with necessity, and that he took analyticity to be a property of sentences. Given his view that claims about which properties an object has essentially or accidentally are claims about which statements about the object are necessary, he naturally concluded that claims about the essential or accidental properties of objects must in the end be claims about which sentences containing terms designating those objects are true in virtue of meaning. Since there is no direct and natural link tying explicitly linguistic claims about the meanings of sentences with equivalent claims about the essential or accidental properties of objects, he naturally concluded that such claims about objects must be confused.<sup>18</sup> Prior to Kripke's clear articulation and defense of a metaphysical conception of necessity—not tied to or dependent upon linguistic concepts like analyticity—there was no way that philosophers concerned with these issues could have seen them clearly. Once the distinction was made, the clarity and utility of the notion of rigid designation and the intelligibility of essentialist claims became inseparable, and all but irresistible.

*Rigid Designation, Possible Worlds, and Criteria of  
"Transworld Identification"*

We next connect the notion of rigidity with the nature of possible worlds, and with another issue, or pseudo-issue, that Kripke takes up in lecture 1—the need for criteria of “transworld identification.” It is sometimes said that before we can evaluate the truth or falsity of the claim that it could have been the case that Nixon, say, was so and so, we have to settle the question of who counts as Nixon in different possible worlds. There are several different ideas connected with this that Kripke rejects. First, if—for some reason (perhaps simply because you are taken in by the *world*-terminology)—you think that possible worlds are large concrete objects—alternative universes that really exist, but do so in a part of space and time inaccessible to us—then it may seem obvious to you that each of us inhabits only one world (just as each of us inhabits only one location at a time on Earth). On this picture, none of the individuals who exist in the other worlds can be identical with any of us, Nixon included. According to this picture, the

<sup>18</sup> The story of this line of reasoning is very well told by John Burgess in “Quinus ab omni naevo vindicatus.” As Burgess points out, it did not help the dialectical situation any that Quine's main pre-Kripkean opponents—who defended *de re* modality, essentialism, and quantifying in to modal constructions—standardly explained their conception of necessity by identifying it with analyticity or logical truth.

most one can do is establish criteria for who in those other worlds is similar enough to our Nixon so that they can appropriately be described as playing the "Nixon role." (As if one could ever learn about such worlds at all.) This is not Kripke's conception of possible worlds.

For Kripke, a possible world is a possible world-state—a way that everything could have been. It is, in effect, a maximal property that the universe could have had.<sup>19</sup> To say that there are possible worlds in which Nixon lost the election is just to say that there are properties that the universe could have had which are such that, if the universe had them, Nixon would have lost the election. In specifying these properties—i.e., these world-states—we can refer directly to Nixon himself. We don't have to come up with descriptive criteria that must be fulfilled if one is to play the "Nixon role."

We may put this another way. In specifying possible world-states, we are not restricted to using general descriptive terms. We are not restricted to saying things like *the world-states are those which are such that had the universe been in any one of them, then someone who graduated from a small college in the most populous state in the country and who was a former vice president, who later became president but was forced to resign, would have been so and so*. If those were the only kinds of descriptions we could give of possible world-states, then we would need descriptive criteria of identity to find out which individual, if any, could be taken to be Nixon with respect to a given world-state. However, we are not restricted in this way. There is no reason why pure descriptions, not containing any names or other rigid designators, should have preferred status in specifying world-states. We can, if we like, specify a class of world-states as those in which Nixon has a certain property P. Since 'Nixon' is a rigid designator, in doing this we are specifying the world-states as those in which a certain individual has P. Given these world-states, we don't, then, have to decide who Nixon is.

Kripke addresses this point in the following passage:

The tendency to demand purely qualitative descriptions of counterfactual situations has many sources. One, perhaps, is the confusion of the epistemological and the metaphysical, between a priority and necessity. If someone identifies necessity with a priority, and thinks that objects are named by means of uniquely identifying

<sup>19</sup> For useful discussions of essentially this conception of possible worlds, see Robert Stalnaker, "Possible Worlds," *Nous* 10 (1976): 65–75, reprinted in Michael Loux, ed., *The Actual and the Possible* (Ithaca, NY: Cornell University Press, 1979); and Nathan Salmon, "On the Logic of What Might Have Been," *The Philosophical Review* 98 (1989): 3–34.

properties, he may think that it is the properties used to identify the object which, being known about it *a priori*, must be used to identify it in all possible worlds, to find out which object is Nixon. As against this, I repeat: (1) Generally, things aren't 'found out' about a counterfactual situation, they are stipulated; (2) possible worlds need not be given purely qualitatively, as if we were looking at them through a telescope.<sup>20</sup>

Although the chief contention of this passage seems, clearly, to be correct, a final point of clarification is in order. When Kripke talks of *stipulating a possible world*, he doesn't mean that which things are possible is a matter of our stipulation. He means that which possibilities we choose to single out, or talk about, is a matter that is up to us. It is up to us to stipulate, or specify, which of the possible states that the world genuinely could have been in we are interested in, and wish to make claims about. Further, the fact that we can use the name 'Nixon' when we specify a class of possible world-states doesn't mean that no such specifications can fail. We might try to specify possible world-states in which Nixon has a certain property P, but fail to do so because, in fact, Nixon couldn't have had that property. For example, we can't successfully stipulate a possible situation in which Nixon is an inanimate object. In this sort of case there is no such possible world-state corresponding to our specification. That is just to say that our specifications are fallible; indeed, the observation that they are fallible shows that facts about possibility are not created or determined by our stipulations. The same point holds for attempted specifications involving descriptions. In general, descriptions have no priority over names or other rigid designators in specifying possible world-states.

### Why Descriptions Normally Do Not Fix the Referents of Names

#### *Arguments against the Weak, Reference-Fixing Version of the Description Theory*

Let us review where we are. So far in our discussion of *Naming and Necessity* we have done the following things:

(i) We defined the notions of rigid and non-rigid designators, and argued that proper names are rigid, whereas most descriptions associated with names by speakers are not.

<sup>20</sup> *Naming and Necessity*, pp. 49–50.

(ii) We discussed Kripke's modal argument that names do not mean the same as non-rigid descriptions that speakers associate with them. In addition, we extended this argument to show that names do not mean the same as descriptions rigidified using the actuality operator.

(iii) We clarified the distinction between essential and nonessential properties, and attempted to explain what Kripke means by talk of possible worlds.

Having done this, we will next examine a weaker theory about the relationship between names and descriptions. In lecture 2, Kripke considers the possibility that descriptions may, as he says, *fix the referents of names*, without giving their meanings. The idea is something like this: Even though proper names don't have meanings in the usual sense, something must be responsible for establishing and maintaining the link between a name and what it refers to. Something must determine what one's utterance of 'Aristotle' refers to. Perhaps descriptions semantically associated with names do this after all. That is, it may be that each proper name is associated with certain descriptions that provide the criteria for determining what it refers to—descriptions that are part of its meaning, and that are mastered by competent speakers, even though they do not provide synonyms for the name.<sup>21</sup> Once the descriptions have fixed what the name designates with respect to the actual world-state, the connection between the name and the referent becomes rigid. Thus, with respect to any possible world-state *w*, the name designates the individual that satisfies the descriptions with respect to the actual world-state—whether or not that individual satisfies those descriptions with respect to *w*. For example, suppose that we have a sentence *Fn*, in which the referent of *n* is semantically fixed by a set *D* of descriptions. On this theory, the truth conditions of the proposition expressed by this sentence can be thought of as being determined as follows. First, it is determined which object *o* uniquely

<sup>21</sup> In saying that the descriptive reference-fixing conditions are semantically associated with the name, that they are part of the meaning of the name, and that they are mastered by competent speakers, we distinguish the interesting, though contentious, claim that the referent of *n* is semantically fixed by a description from the trivial and uninteresting claim that it is possible to describe the process by which *n* got its referent. For any word whatsoever—and, necessarily, if, therefore, obviously, Aristotle, 3, etc.—there is some description that correctly describes the process by which the word acquired its meaning, or its reference. However, that does not mean that all words have their meanings, or referents, semantically fixed by descriptions. It is one thing to describe how words got to have the meanings and referents they do; it is another to say that descriptions are parts of the meanings of words in the manner contemplated by the description theory.



satisfies *D* with respect to the actual state of the world. Then, given this object, we can see that the proposition expressed by the sentence is true with respect to an arbitrary possible world-state *w* iff *Fx* is true of *o* relative to *w*.

In the process of testing this theory, Kripke isolates corollaries (i–iv) of the theory.

The referent of a name *n* (for a speaker at a time) is semantically fixed by a description, a conjunction of descriptions, or a cluster of descriptions *D*. If *D* fixes the referent of *n*, then:

- (i) the speaker believes that *D* applies to a unique individual;
- (ii) if *D* does apply to a unique individual *o*, then *o* is the referent of *n*;
- (iii) if *D* does not apply to a unique individual, then *n* has no referent;
- (iv) the speaker knows (or is capable of knowing) *a priori* that the sentence *If n exists, then n is D* expresses a truth.

In lecture 2, Kripke goes through these theses one by one and gives what he takes to be counterexamples to them. I won't belabor this, but will try to say enough to indicate the main idea. In so doing, I will employ a strategy that is implicit in Kripke's discussion. I will take it that the descriptions that are candidates for fixing the reference of a name *n* for a speaker are more or less those that the speaker would give, either initially or subject to some reasonable idealizations, if asked *Who or what do you use n to refer to?* So, in giving these counterexamples, I will assume that if a speaker uses descriptions to semantically fix the referent of a name, then he would normally either already be aware, or easily be capable of becoming aware, of those descriptions, and be able to supply them if asked.<sup>22</sup> Later, when we get to Kripke's own positive theory about how reference is determined, we

<sup>22</sup> Some post-Kripkean descriptivists do not accept this limitation on candidates for descriptions that semantically fix referents. My own view is that once this limitation is abandoned, it becomes difficult to make the important distinction indicated in the previous footnote between descriptions of the pre-semantic causal process by which words acquire meaning and reference, and descriptions which are part of the meanings of terms that are mastered by competent speakers. For a critique of recent versions of descriptivism that, in my view, run afoul of this distinction, see the works cited in footnote 10.

will consider the question of whether, by relaxing this condition, his own positive theory could be put in the form of a description associated with the name by a speaker.

We now consider the corollaries. First, corollary (i). Kripke observes that in the case of many names, the descriptive information speakers associate with the name is too impoverished to pick out an individual uniquely. One example of this is provided by the name *Cicero*. What do most of us know about Cicero? Many don't know much more than that he was a famous Roman, perhaps a statesman and orator of some sort. Presumably, however, there was more than one famous Roman statesman and orator. Most of us recognize this. So, Kripke would say, we don't even believe that the description we associate with the name picks out one individual uniquely. Hence, this is a counterexample to corollary (i). Nevertheless, our use of the name *Cicero* does refer to a unique person. Kripke concludes from this that the linguistic mechanism that determines the referent of our use of the name must be something other than that maintained by the descriptive theory.

There is a point here that is worth noting, which shows that this sort of example is more common than one might first think. Imagine that we have a speaker who knows more about Cicero than that he was a famous Roman statesman and orator. Suppose he knows a certain fact that Quine is fond of stating—namely, that Cicero was the famous Roman statesman who first publicly denounced Catiline. Now this description really does pick out the man Cicero uniquely. So you might think that, for this particular speaker, Kripke's objection to corollary (i) doesn't work. However, one should be cautious about drawing this conclusion. For the description in question itself contains a proper name—*Catiline*. And one might ask the speaker to give his description of Catiline. If the speaker is like many of us, the best he could do would be to say that Catiline is the Roman leader first publicly denounced by Cicero. So what we have is a pair of names, *Cicero* and *Catiline*, each associated with a description that determines a unique individual, but only if the name contained within the description already has a reference independently. If the speaker's information is exhausted by these descriptions, then the description theory will not be able to explain how the referent of either name is determined.

This example brings out a strong requirement imposed by the pure form of the description theory, when it is taken to be a theory about how the referents of all names are fixed. What the theory requires is that each name be associated with **purely descriptive properties** that

are sufficient to uniquely determine its referent. But this is highly counterintuitive. If you ask yourself whether you have such properties associated with every name you use, you will, I am confident, come to the conclusion that you don't. If this is right, then speakers do not even take themselves to have the sort of descriptive information associated with each name that the pure form of the theory requires. So much for corollary (i).

On to corollaries (ii) and (iii). The *Cicero*-type examples we have been discussing are cases in which we refer to an individual, even though the purely descriptive information at our disposal fails to pick out any individual uniquely. Another kind of case Kripke considers is one in which the problem is not a lack of information, but rather the existence of misinformation. Consider the name *Thales*. About all that I know about Thales is that he was a pre-Socratic philosopher who held that all is water. Suppose, however, that there was a certain individual called *Thales* by his contemporaries, or at any rate called by some name which when translated and passed on to us has come down as *Thales*. Suppose further that the man's contemporaries attributed a view to him that he never held. Suppose he never held that all was water, but rather believed something more sensible. Nevertheless, the story about him spread and changed, and all that has come down to us now is that Thales held that all is water. In this imaginary case, the description we associate with the name doesn't designate the person the name really refers to. It could even be that there was some other pre-Socratic philosopher who was a hermit no one knew about. Even if by some chance he really did hold that all was water, and so satisfied the description we associate with the name, that would not make him Thales. The name we use would not refer to him, but rather to the originally misunderstood philosopher. These points are *prima facie* indications that corollaries (ii) and (iii) of the description theory are false.

Another case of this type that Kripke cites is that of Peano. The main thing that most people who have heard of Peano believe about him is that he was the originator of the now standard axioms of elementary arithmetic—the so-called *Peano axioms*. In fact, he did publish those axioms, and people remember him for it. However, in a footnote he credited the axioms to another mathematician—Dedekind. The footnote was mostly forgotten, and Peano ended up being credited by most people with the axioms. Let us take it that the axioms really did originate with Dedekind. Before I read *Naming and Necessity* I don't think I knew that, though I had heard of the *Peano axioms*. If I had

any description associated with the name *Peano* at that time, it was, I think, *the discoverer of the axioms of arithmetic*. Still, my use of the name *Peano* did not refer to Dedekind.

Perhaps, though, it will be objected that I had some other description associated with the name that really did refer to Peano. Let's consider some candidates for such a description. The first candidate is the parasitic description:

*the person to whom most people refer when they use the name 'Peano'*

The idea behind this suggestion is that a person who doesn't know enough to describe Peano uniquely on his own can fix the referent of his use of the name via this parasitic description, provided that most other people themselves have the resources to refer correctly and uniquely to Peano. The problem with this idea is that it runs the risk either of failure or of circularity. If most people don't have other descriptions that independently succeed in referring to Peano, then the description theorist is stuck with the result that those who use our parasitic description fail to refer. But we can imagine cases in which most people lack such non-parasitic descriptions; in fact this may still be true in the case of the name *Peano*. Nevertheless, people in these situations do refer to someone, and the person they refer to is Peano.

What about other candidates for the reference-fixing description? Consider:

*the person to whom most experts refer when they use the name 'Peano'*

This is no good, since it doesn't specify what sort of experts—in Italian opera? Renaissance painting? etc.—we are talking about. Do we do any better with *the person to whom most mathematicians refer when they use the name 'Peano'*? Well, mathematics is a large field, and it could easily happen that most mathematicians have only the Dedekind-description associated with the name. Nevertheless, they still use the name to refer to Peano, not Dedekind. How about the description *the person to whom most Peano-experts refer when they use the name 'Peano'*? But this will only take us in a circle, since in order to find out who Peano is, we have first to locate the Peano experts, whereas to find out who is a Peano-expert—i.e., who has expert knowledge of the man Peano—we first must find out who Peano is.

A similar circularity affects

*the person to whom the axioms of arithmetic are standardly attributed*

What is it to attribute the axioms of arithmetic to someone? I suppose it is to say of that person that he discovered them. Well, what would people say if we asked them, *Who discovered the axioms of arithmetic?* A lot would probably answer *Peano*, thereby saying of whomever that name refers to that he discovered the arithmetical axioms. So who does the name refer to? According to the theory, it refers to the person who satisfies the description *the person to whom the axioms of arithmetic are standardly attributed*. Thus, in order to determine the person who satisfies the description, we first have to get the referent of the name *Peano*, but, if this version of the description theory is correct, we can't do that until we first determine who satisfies the description. Hence, we have gone in a circle again. The upshot of all of this is that it is hard to come up with any non-circular, reference-fixing description that is not subject to clear and obvious counterexamples. On the basis of this, Kripke concludes that corollaries (ii) and (iii) of the reference-fixing version of the description theory are false.

This brings us to the final corollary of the description theory.

- (iv) The speaker knows (or is capable of knowing) a priori that *if  $n$  exists (existed), then  $n$  is (was)  $D$*  expresses a truth.

Why is this corollary part of the theory? Well, if the linguistic rule by which the referent of a name is fixed is that the name is to refer to whoever or whatever is designated by a certain specific description  $D$ , then one knows (or is capable of knowing), simply by virtue of knowing this rule, that the sentence *If  $n$  exists (existed), then  $n$  is (was)  $D$*  cannot fail to be true. Why? Because if  $D$  fails to designate anything, then  $n$  doesn't either, and the sentence is trivially true by virtue of the falsity of its antecedent. On the other hand, if  $D$  does designate something, then  $n$  designates the same thing, and the consequent is guaranteed to be true. Either way, the conditional sentence as a whole is guaranteed to be true.<sup>23</sup> All of this one knows (or is capable of knowing) just by knowing the rules of the language, without doing any extra empirical investigation. Though one might certainly raise the question of whether this is what genuine apriori knowledge really amounts to, for the moment we will follow Kripke's lead in discussing his examples

<sup>23</sup> If one is worried that a sentence containing a name that fails to refer may not express a proposition, and so fail to be true, then one may change the statement of the corollary to read, *The speaker knows (or is capable of knowing) a priori that if ' $n$  exists' expresses a truth, then ' $n$  is (was)  $D$ ' also expresses a truth*. Since this issue does not affect the outcome of the present discussion, I will leave it to one side.

and mean by *apriori* nothing more than that which can be known solely on the basis of understanding one's language.<sup>24</sup> Hence, if the reference-fixing version of the description theory is correct, corollary (iv) should always hold.

However, when we consider particular names, we see that it typically fails. Consider the name *Columbus*, for example. Here, the most salient description that many people associate with the name is something like *the European who sailed from Spain in 1492 looking for a new route to Asia, but ended up discovering America instead*. However, the linguistic rule determining the referent of the name is surely **not** that it is, by definition, whoever satisfies this description. For if we ask ourselves how we know that if Columbus really existed, then he was the European who sailed from Spain in 1492 looking for a new route to Asia, but ended up discovering America instead, we surely will **not** say that we know this apriori, simply by understanding the language. Rather, we know this because we read it in textbooks, and because our teachers told us. And how did our sources know this? Presumably by similar means, involving various kinds of scholarship, testimony, surviving documents, artifacts, and the like. But if that is so, then our knowledge rests on empirical evidence, and so is aposteriori. Moreover, our belief about Columbus remains subject to revision, if some of the historical evidence we rely on is shown to be fake, faulty, inaccurate, or seriously incomplete. Although we don't expect it to happen, it is certainly conceivable that new evidence might be discovered showing that Columbus never left Spain, but sent someone else in his place. This shows that neither the proposition expressed by (15), nor the claim that sentence (15) expresses a truth, is knowable apriori in the sense we have indicated.

15. If Columbus existed, then Columbus was the European who sailed from Spain in 1492 looking for a new route to Asia, but ended up discovering America instead.

Thus it looks like all the major corollaries of the reference-fixing version of the description theory are false—when taken to express universal generalizations covering all proper names. On the basis of this, Kripke concludes that **there is no general semantic rule** specifying that the referent of a name is the individual designated by descriptions associated with it by speakers. However, this does not mean that we

<sup>24</sup> This conception of the apriori will be revisited and critically evaluated in chapter 16.

never fix the referent of a name descriptively. On the contrary, Kripke thinks we can and, in rare cases, do semantically fix the referent of a name by stipulating that it is to be whatever satisfies a certain description. He gives the example of the name *Neptune*, which he suggests may originally have been introduced as a proper name for whatever caused certain perturbations in the expected orbit of the planet Uranus. Kripke insists that even if the name is introduced by a stipulation that it is to refer to whatever turns out to satisfy a certain description, still the name is **not** synonymous with the description, but rather is a rigid designator. Moreover, even if the name is first introduced by a reference-fixing description, later, when it is passed from speaker to speaker, the description may be lost and the name may come to be understood without reference to the description. If Kripke is right, such may have been the history of the name *Neptune*. Surely, however, it is no longer apriori for us that the sentence *If Neptune existed, then Neptune caused perturbations in the expected orbit of Uranus* expresses a truth (if it ever was).

*Kripke's Historical Chain Conception of  
Reference Determination*

So according to Kripke, the reference of a name is occasionally, semantically fixed by a description. But in most cases that is not how the reference is initially determined; and even when it is, the semantic association of the name with the reference-fixing description is likely to be short-lived. This raises the question of how, in the vast majority of cases, the reference of a name is determined. Kripke presents a positive picture of reference-determination that attempts to answer this question. His idea is extremely simple and commonsensical. A particular name is introduced for an object or person. After the name has been introduced by some sort of linguistic baptism, those introducing the name begin to use it in conversation to refer to its bearer. New people hear the name, and begin to use it themselves, intending to refer to the same individual that their sources used it to refer to. This process continues, with the name being passed from one user to the next, with each user forming a link in a chain of reference transmission, or inheritance. Often, some descriptive content will accompany the passing of the name, but typically this content will vary from speaker to speaker, and as the chain of use gets longer, it may end up that for many speakers the name has more misinformation than accurate information

about its referent associated with it. No matter. The information that a person ends up associating with the name is **not** what determines its referent. Rather, the referent of a use of a name by a speaker  $x$  is determined by the historical chain connecting  $x$ 's use to speakers from whom  $x$  acquired the name, connecting those speakers to their sources, and leading ultimately back to the individual baptized with the name. According to this picture, if a person picking up a name intends to use it to refer to whatever individual is referred to by his sources for the name, then it usually doesn't matter very much which other beliefs he may have about the referent. What determines reference is not a speaker's beliefs, but the chain of use in which the speaker stands. For Kripke, referring is typically not something that one does in isolation; it is a community effort.

That is the general picture. It should be noticed, however, that Kripke is not explicit about the exact nature of the chain of use, or about just which facts have to hold in order for a use of a name to stand as one link in a chain that determines later reference. Suppose, for example, that the person from whom I first picked up the name *Plato* was talking about his neighbor, whom he believed to be very wise. Suppose that after speaking to this person I had many other conversations in which the name *Plato* was used to describe Socrates' famous biographer. At a certain point, I read about Plato, and so discovered the name in print. I even read translations of Plato's work. All of this could be true, even if I wrongly assumed that the person from whom I first heard the name was talking about the same individual as everyone else. In this sort of case, to whom do I now refer when I use the name *Plato*? Do I refer to the wise neighbor of my ultimate source, as he himself did? Or do I refer to the ancient philosopher, as do most of my other sources for the name? Surely, the latter answer is correct. However, Kripke nowhere presents a precise and explicit theory which gives a clear verdict about potentially problematic cases like this. His aim, he says, is not to provide such a theory, but rather to sketch an alternative picture of how reference standardly works—a picture in which a historical chain of uses **somehow** links the speaker to the referent, rather than a picture in which the link between speaker and referent is provided by descriptions the speaker associates with the name.

The incompleteness of Kripke's picture is also illustrated by the phenomenon of reference change. Sometimes a name can start off referring to one thing, then go through a period in which that thing is confused with something else, and still later it will count as referring only to the



second thing—even if no one ever consciously tried to redefine the name, or change its reference. Such a case is presented by Gareth Evans in an article called “The Causal Theory of Names.”<sup>25</sup> He gives the example of the name *Madagascar*. He maintains that a version of the name was originally used to designate some part of the mainland of Africa. However, when the Arabs and Europeans arrived, they picked up the name from the inhabitants, mistakenly thinking it was a name of the big island off the southeastern coast of Africa. As Evans tells the story, they used it intending both to refer to the island and to refer to the area that the locals used it to designate, thinking that the two were one and the same. After a period of time, the referent of the name became, unambiguously, the island. If genuine, examples of this sort do not discredit Kripke’s historical conception of how the referent of a name is fixed. However, they do show that there is a substantive, nontrivial question of what goes into establishing the links in the historical chain.

*Does the Historical Chain Theory Provide Descriptivists  
with Reference-Fixing Descriptions?*

With this in mind, we may turn to the suggestion sometimes made by die-hard descriptivists that what Kripke has done is simply to give a specific kind of reference-fixing version of the description theory of names. Their idea, in simplest terms, is that the reference of a name *n* for a particular speaker is semantically determined by some description extractable from Kripke’s historical chain theory of reference transmission and inheritance. David Lewis states the idea as follows in his 1997 paper, “Naming the Colours”:

Did not Kripke and his allies refute the description theory of reference, at least for names of people and places? Then why should we expect descriptivism to work any better for names of colours and colour experiences? . . . I disagree. What was well and truly refuted was a version of descriptivism in which the descriptive senses were supposed to be a matter of famous deeds and other distinctive peculiarities. A better version survives the attack: *causal descriptivism*. The descriptive sense associated with a name might for instance be ‘the place I have heard of under the name “Taromeo”’ or maybe

<sup>25</sup> *Proceedings of the Aristotelian Society*, supplementary volume 47 (1973), 187–208; reprinted in his *Collected Papers* (Oxford: Clarendon Press, 1985).

‘the causal source of this token: Taromeo’, and for an account of the relation being invoked here, just consult the writings of causal theorists of reference.<sup>26</sup>

To keep things simple, we may illustrate this idea using the description *the individual to which the person or persons from whom I acquired the name referred when they used it*. One of the things that makes the idea that a person’s use of a name is fixed by such a description seem plausible is the requirement, recognized by Kripke, that in order for a chain of reference transmission to be created by passing a name from one speaker to the next, the person acquiring the name must intend his reference to be parasitic on the reference of his sources. The descriptivist may be seen as proposing to put this requirement in the form of a description that semantically fixes the referent of the name.

Although the idea may appear reasonable, there are, I think, several reasons to be cautious. First, it is not clear that speakers invariably have implicitly in mind, among all the different descriptions they associate with a given name, some precise reference-fixing description for it. We know that the description *the individual to which the person or persons from whom I first acquired the name referred when they used it* does not always pick out my referent for *n*. Moreover, it is not fully clear even to theorists precisely which parasitic description of this sort will be sufficient to handle all the different problematic cases. At present there simply is no precise and explicit historical chain theory that one can turn into a description that is adequate for every case. Moreover, even if theorists were to come up with such a description, it is far from obvious that ordinary speakers must have had it implicitly at their disposal all along, whenever they used a proper name.

Second, even if we assume that speakers always have an appropriate parasitic description associated with a name, it would have to be shown that they somehow accord this description priority in determining the referent of the name over all other descriptions that they associate with it. This might not be easy. If we were to ask ordinary speakers to give us the descriptions that most reliably specify the referents of the different names they use, it is clear that they would **not** spontaneously volunteer the relevant, non-circular, parasitic descriptions.

<sup>26</sup> P. 353, footnote 22 of “Naming the Colours,” in his *Papers in Metaphysics and Epistemology* (Cambridge University Press, 1999), originally published in *Australasian Journal of Philosophy* 75 (1997). Other contemporary descriptivists express similar ideas. For example, see David Chalmers, “On Sense and Intension,” *Philosophical Perspectives* 16, *Language and Mind* (Oxford: Blackwell, 2002), 135–82.

Perhaps, if we were to guide them through enough Kripke-style thought experiments about reference, we would have some success in eliciting parasitic descriptions that approximate correct results for many cases. The descriptivist might then postulate that, like the slave boy in Plato's *Meno*, these ordinary speakers must have had, and unconsciously given priority to, the relevant reference-fixing descriptions all along. Surely, however, this story must be regarded as speculative in the extreme.

Third, throughout all of this, it is important to keep in mind the distinction mentioned in footnotes 21 and 22 between descriptions of the pre-semantic causal processes by which words acquire and retain their meanings and/or referents, and descriptions which are parts of the meanings thereby acquired, and hence grasped by competent speakers who have mastered the words. Although there are clearly descriptions of the first sort for every word in our language—including *if*, *and*, and *but*—this does not show that there are descriptions of the second type for all of them. The doctrine that the referents of names are semantically fixed by descriptions associated with them by speakers is an ambitious and problematic thesis that asserts the existence of descriptions of the second type for all names.

Finally, one shouldn't think that the reference-fixing version of the description theory—thought of as a semantic theory implicitly mastered by competent speakers—somehow **has** to be correct. It could be that there is a certain process by which reference is passed from speaker to speaker. It could also be that, standardly, when one picks up a name, one intends to use it to refer to whatever the person one picked it up from uses it to refer to, or to whatever it is properly used to refer to in one's linguistic community as a whole. Once one has acquired the name, one can start using it to express beliefs about the object it stands for. Later, one may forget completely about all aspects of how one acquired the name, while retaining the relevant beliefs one uses the name to express. If, at this later time, one asks what determines the referent of one's use of the name, then the answer might be that it simply inherits its reference from the beliefs one has been using it to express—whether or not one can come up with any correct description about how one originally acquired the name, or any correct description of one's later uses of the name that may have changed its original reference. (Think of the Plato case.) If this is right, then there may be a natural process of reference inheritance by which later uses of a name inherit their reference from earlier uses, even if speakers themselves

don't have to have any complete and accurate understanding of the process. There is nothing incoherent about this idea. So there is nothing outlandish about the idea that there might be no correct reference-fixing version of the description theory that is part of what one learns when one masters a language.<sup>27</sup>

Finally, there is, I think, an unclarity in *Naming and Necessity* about the kind of question we are asking when we ask how the reference of a name is fixed, for a speaker at a time. Often, Kripke writes as if this were a semantic question about a rule that speakers master when they learn a name—a rule of the sort “*n*” as used in context *C* refers to (a) whatever satisfies such and such descriptions, or (b) whatever stands at the end of such and such chain of reference transmission. On this way of looking at things, the question *How is the referent of a name fixed?* is used to express is a request for a semantic rule which is implicitly grasped by speakers, and which could be used by them in specifying the truth conditions of sentences containing the name. But this is not the only type of question that the sentence could be used to asked.

Sometimes, for example, it may be used to ask a pragmatic question about which of its various contents a term is used with on a particular occasion. For example, one might wonder what determines whether, on a certain occasion, ‘David’ is used to refer to David Kaplan or to David Lewis, just as one might wonder what determines whether, on a particular occasion, the word ‘bank’ is used to talk about a river bank or a financial institution. The causal sources of these particular uses of the words may very well be relevant to answering these questions. However, the questions are not semantic ones about which meanings are assigned to words; rather, they are questions about how, when a word carries multiple meanings, it is determined which meaning a speaker is using it with.

There is still another type of question that *How is the referent of a name fixed?* may be used to ask. As theorists, we may want to know (i) how a

<sup>27</sup> It should also be remembered that even if some weak version of the description theory were correct, according to which the referents of names were semantically fixed by Lewis-style descriptions, the theory would provide no help in solving the problem that descriptivists have traditionally been most concerned with—namely, specifying the content of the propositions expressed by sentences containing names, and believed and asserted by speakers who accept, or assertively utter, those sentences. When I say that the ancient Babylonians believed that Venus was a star, I am not saying that they believed that the object which is connected by an appropriate historical chain of use to my utterance of the word *Venus* was a star. So even if, as seems extremely unlikely, some such version of descriptivism could be made to work as a semantic account of reference-fixing, it would **not** solve the problem of content that historically has been of most concern to descriptivists.

name, or any other expression, initially came to have the meaning and/or reference it does and (ii) in virtue of what aspects of its use by speakers it retains that meaning and reference in the common language of the community. These are theoretical questions that may arise even after it has been settled precisely what the meaning and reference of a given expression is, either in the language of the community as a whole, or as used by a speaker on a particular occasion. These questions are foundational questions about the causal processes that originally endowed expressions with their semantic properties, and that maintain them in the language with the properties they have.

When Kripke discusses the fix-the-referent version of the **description theory**, he is clearly discussing a proposal about the semantics of names. Because of this, his way of framing the general discussion—*What fixes reference?*—may seem to suggest that the two answers he gives to this question—*Descriptions in certain unusual cases* and *Historical chains of reference transmission for the great mass of proper names*—are on a par. Since the claim about descriptions clearly may be understood to be semantic, and hence about the linguistic rules that speakers must master in learning various names, this encourages the unwary reader wrongly to think of the claim about historical chains in the same way. This, I believe, is the source in Kripke's discussion of the descriptivists' last-gasp attempt to interpret his historical chain theory of reference transmission as providing the descriptions needed for a correct descriptive theory of the semantics of names. This is what leads them to their view that, details aside, the reference of *n* for a particular speaker is semantically determined by a description such as *the individual to which my sources referred when they used n*.

In my opinion, this whole way of thinking is wrongheaded. There is nothing special here about the semantics of names. Standardly, when one uses any word in the language of one's community, one does so with the intention that it should carry whatever meaning and reference it has already acquired. This is a fact about the use of all expressions, not about the semantics of any of them. To the extent that there are additional questions about names, they are pragmatic and foundational questions. Kripke's comments about initial baptisms should be seen as answers to the foundational question *In virtue of what did these terms originally come to refer in the language to what they do?* His comments about causal chains of reference transmission should be understood as providing information relevant to answering the pragmatic question *How is it determined which of many bearers of the name a particular*

*utterance of it refers to?* and the foundational question *In virtue of what does the name continue to refer in the language to this object?*<sup>28</sup> If one then asks, *And what should a semantic theory of such names tell us?*, the most reasonable answer, I believe, is that it should tell us what these terms refer to in the common language of the community, and nothing more. However, this takes us beyond what Kripke explicitly committed himself to, and to the brink of one of the fundamental unsolved problems raised by *Naming and Necessity*. This problem, which is centrally relevant to Kripke's groundbreaking discussion of the distinction between necessary truth and apriori truth, is one we will take up in the next chapter.

<sup>28</sup> This way of looking at Kripke's discussion of reference-fixing via baptism and historical chains of reference transmission is explained and forcefully defended by my former student Jonathan McKeown-Green in his unpublished Princeton doctoral dissertation, *The Primacy of Public Language* (2002).

# CHAPTER 15

## THE NECESSARY APOSTERIORI

### CHAPTER OUTLINE

1. *A Framework for Discussing the Necessary Aposteriori and the Contingent Apriori*

Propositions as contents of sentences, objects of belief, and bearers of truth

2. *Genuine Instances of the Necessary Aposteriori: Explanation and Consequences for Our Conception of Inquiry*

Inquiry as locating the actual world-state within the space of metaphysically possible world-states; the incompatibility of this conception with the necessary aposteriori. Why Kripke's explanation of genuine instances of the necessary apriori shows that the actual is sometimes epistemologically prior to the possible

3. *Identity Statements and the Necessary Aposteriori*

The gap in Kripke's arguments that true identity statements involving names are often knowable only aposteriori, and a plausible reconstruction of his reasoning using disquotational principles relating acceptance of sentences to belief in the propositions they express; the apparent falsity of strong disquotational principles; the ineffectiveness of plausible weak disquotational principles; the source of these problems in the nontransparency of meaning; final assessment: Kripke's arguments concerning names present us with a dilemma

4. *A Further Note on Inquiry*

### A Framework for Our Discussion

The claim that there are genuine examples of the necessary aposteriori, and the corresponding claim that there are genuine examples of contingent apriori, are among the most important and far-reaching doctrines of *Naming and Necessity*. In this chapter, I will explain and evaluate the first of these claims, and the arguments that Kripke gives for it. In chapter 16, I will do the same for the second claim. Kripke's discussion of these topics may properly be regarded as ground-breaking, and I

will argue that many of his examples and arguments are illuminating. However, I will also indicate ways in which I believe he muddied the waters by saying certain things that are puzzling at best, and defective at worst. In the case of the necessary aposteriori, much of what he says can, I think, be taken to be straightforwardly correct, though, as we will see both in this chapter and in chapter 17, certain parts of his discussion unnecessarily weaken his case. In the case of the contingent apriori, considerably more reconstruction and revision is needed. Our aim in this chapter and the next will be to clarify and correct these matters, in order to arrive at clear and defensible versions of the two doctrines.

In bringing out both the puzzling and the illuminating aspects of Kripke's discussion, I will, in these two chapters, make use of a modest theoretical framework that goes beyond what he explicitly commits himself to in *Naming and Necessity*. The central assumptions of this expository framework are the following:

- A1. Some things are asserted, believed, and known. Propositional attitudes like assertion, belief, and knowledge are relations that hold between agents and the things that they assert, believe, and know.
- A2. The things asserted, believed, and known may be expressed by sentences, and designated by clauses such as *the statement that S*, *the assertion that S*, *the belief that S*, *the claim that S*, *the proposition that S*, or simply *that S*. I will call the things designated by these clauses *propositions*.
- A3. Propositions are bearers of (contingent or necessary) truth and falsity.
- A4. Propositions are not identical with sentences used to express them. Whatever they turn out to be, propositions are, roughly speaking, things which different sentences that "say" or express the same thing have in common.
- A5. Propositional attitude ascriptions—*x asserts/believes/knows/knows apriori/knows aposteriori that S*—report that an agent asserts, believes, knows, knows apriori, or knows aposteriori the proposition designated by *that S*.

Kripke's discussion of the necessary aposteriori and the contingent apriori indicates his commitment to the view that (i) there are genuine



cases in which some one thing is both necessarily true and knowable only *aposteriori* (on the basis of empirical evidence), and (ii) there are genuine cases in which some one thing is both contingent and knowable *apriori* (without appeal to such evidence). From the point of view of our modest theoretical framework, this is the view that there are some propositions—i.e., things capable of being asserted, believed, and known—that are both necessarily true and knowable only *aposteriori*, and there are other propositions that are both contingently true and knowable *apriori*. Although I will argue that this view is correct, I will also raise serious questions about Kripke's treatment of a substantial range of cases.

### **Genuine Instances of the Necessary *Aposteriori*: Explanation and Consequences for Our Conception of Inquiry**

There is a natural and initially attractive conception of inquiry according to which ignorance about a given subject is a matter of lacking information about which, of certain relevantly different possible states the world could be in, it is actually in; and complete ignorance is a condition in which one doesn't know which, of all the possible states that the world could be in, it is actually in. According to this conception, when an agent is in this condition, (i) all metaphysically possible states of the world are epistemically possible—i.e., every way that the world could possibly be is a way that, for all the agent knows, it might actually be, and (ii) every epistemic possibility is a metaphysical possibility—i.e., every way that, for all the agent knows, the world might be is a way that the world really could be. Inquiry is the process of escaping from this position of ignorance. By investigating the world or relying on the testimony of others, the agent learns contingent truths that distinguish the way the world actually is from other ways it might possibly be, but isn't. Each time the agent learns one of these truths, he narrows down the class of metaphysical/epistemic possibilities compatible with what he knows, and within which he locates the way the world actually is. According to this conception, acquiring information is equated with narrowing down the range of metaphysically possible world-states that are compatible with what one knows. We may also speak of the truth of one proposition as providing information supporting the truth of another. On this conception, the truth of a proposition

p provides information supporting the truth of a proposition q by ruling out certain possible ways in which q might fail to be true. Thus, the truth of p supports the truth of q only if the set of possible world-states with respect to which both p and the negation of q are true is non-empty.

There are two immediate consequences of this conception of inquiry. The first is that necessary truths are uninformative. Since they are true with respect to all possible world-states, knowledge of them provides no information, and is irrelevant to locating the way the world actually is within the range of possible ways it might be. Second, there are no necessary truths which, though knowable, are knowable only a posteriori. To say that a proposition q is knowable only a posteriori is to say that one can have the justification required to know q only if one has empirical evidence supporting its truth. However, according to the conception of inquiry just sketched, this is impossible. In order for the truth of any proposition p to support the truth of q, and hence to provide evidence for it, there must be possible world-states with respect to which q is untrue, which are ruled out by the truth of p. Since q is necessary, there are no possible world-states with respect to which it is untrue; hence there can be no evidence for q. This means that, on the conception of inquiry just sketched, there can be no necessary truths which, though knowable, are knowable only a posteriori.

Although a number of philosophers have taken this conception of inquiry, and the consequences that follow from it, to be plausible and even axiomatic, the conception is directly challenged by the framework developed by Kripke in *Naming and Necessity*.<sup>1</sup> This challenge is illustrated by the following examples.

1. Gregory Soames is not identical with (i.e., is not the same individual as) Brian Soames.
2. If Saul Kripke exists, then Saul Kripke is a human being.
3. This table is not made out of clay.
4. If this table exists, then this table is made of molecules.

It seems evident that each of these sentences expresses a proposition that is knowable only a posteriori, on the basis of some sort of empirical

<sup>1</sup> A good example of a philosopher who accepts this conception of inquiry, and extends it to provide a model of discourse, is Robert Stalnaker. See his "Assertion," *Syntax and Semantics* 9 (1978): 315–32, reprinted in his *Context and Content* (New York: Oxford University Press, 1999), and his *Inquiry* (Cambridge, MA: MIT Press, 1984).

evidence. In the case of (1), one needs to find out who Gregory and Brian are, and to assure oneself that they are different. In the case of (2), if the question were ever raised as to whether Kripke was a sophisticated robot, or an alien sent from another world, one would need empirical evidence to rule out these possibilities—though, of course, given their fanciful nature, not much evidence would be required. A similar point is true of (3), the justification of which might be provided by a cursory examination of the table. In the case of (4), the evidence required to know the truth that it expresses is much greater, and more sophisticated. Still, since in all four cases empirical evidence is required to know the truths expressed, all four propositions are knowable only *a posteriori*.

They are also necessary. In each case, the subject expression is a rigid designator—the names *Gregory Soames* and *Saul Kripke*, plus the demonstrative phrase *this table*.<sup>2</sup> Because of this, the sentences express necessary truths iff the properties they attribute to the referents of their subjects are essential properties—the properties of being non-identical with Brian Soames, of being human, of being not made out of clay, and of being made of molecules. These do seem to be essential properties; in fact they seem to be essential properties of anything that has them. For example, it is plausible to think that any individual who really is not the same individual as Brian Soames could not have existed while being the very same individual as Brian Soames. Because the name *Brian Soames* is itself a rigid designator, we can also make the point in another way: since the property of being non-identical is an essential property of any pair of things that have it, if two individuals (such as my two sons) really are non-identical, then there is no possible circumstance in which they are one and the same individual. Similar points hold for the other properties mentioned in these examples—anything that really is human could not have existed without being human, any object not made out of clay could not have existed while being (originally and entirely) made out of clay, and anything that really is made up of

<sup>2</sup> To say that an indexical phrase, like *this table*, is a rigid designator is to say, roughly, that when one uses it in context of utterance to refer to a particular thing, that thing remains its referent with respect to all possible world-states with respect to which we might evaluate a sentence containing it. So if I say, *This table is made of wood*, referring to the table *t* now directly in front of me, then I say something (express some proposition) which is true, when taken as a description of any possible world-state *w*, iff had the world been in state *w*, *t* (that very object) would have been made of wood. See David Kaplan, “Demonstratives,” in Joseph Almog, John Perry, and Howard Wettstein, eds., *Themes From Kaplan* (New York and Oxford: Oxford University Press, 1989).

molecules could not have existed without being made up of molecules. Thus, sentences (1–4) all express necessary truths. Since they are also knowable only *a posteriori*, they are examples of the necessary *a posteriori*.

How can this be? How can a proposition that is necessary (and known to be necessary) be knowable only *a posteriori*? Kripke's answer appeals to our knowledge of which properties are essential. He argues, quite plausibly, that we know *a priori* that properties like non-identity, being human, being not made out of clay, and being made out of molecules are essential properties of the things that have them. So we know *a priori* that if things have these properties, then they have them necessarily. This means that the propositions expressed by (1–4) are such that we know *a priori* that **if they are true, then they are necessarily true**. Still, finding out that they are, in fact, true requires empirical investigation. This means that sometimes, in order to find out whether certain things are true with respect to all possible states of the world, and other things are true with respect to no possible states of the world, we first must find out what is true with respect to the actual state of the world. **Sometimes in order to find out what could and could not be, one first must find out what is.** This insight is incompatible with acceptance of the conception of inquiry sketched above. On that conception, necessary truths are uninformative, epistemic possibility is restricted to metaphysical possibility, and all inquiry is a matter of narrowing down the range of (metaphysically) possible states that the world could genuinely be in. Kripke's arguments may plausibly be taken to show that these views, and the conception on which they are based, are simply mistaken.

The central problem with the conception is its restriction of ways things could **conceivably** be to ways things could **really** be—i.e., its restriction of epistemic possibility to metaphysical possibility. This is something that Kripke rejects. Instead of identifying these two kinds of possibility, he sharply distinguishes them.<sup>3</sup> Once this is done, and both rigid designation and the existence of nontrivial essential properties of objects are accepted, the necessary *a posteriori* follows unproblematically.<sup>4</sup>

<sup>3</sup> As we will see in chapter 17, this point requires further discussion. Near the end of lecture 3 of *Naming and Necessity*, Kripke responds to an objection to his conception of the necessary *a posteriori* in a way that may seem to cast doubt on his commitment to a sharp and robust distinction between epistemic and metaphysical possibility. In the end, I will argue that the passage is both misleading and something of an anomaly. However, since it occurs in his discussion of natural kind terms, we will wait until then to analyze it.

<sup>4</sup> Essentially this explanation of the necessary *a posteriori* is given by Kripke on pp. 151–53 of "Identity and Necessity," in Milton Munitz, ed. (New York: NYU Press, 1971).

How, then, should we think of inquiry? Is there any way of modifying the original conception that retains its attractive features, while avoiding its errors? Although Kripke doesn't address this question explicitly, there is a natural strategy for doing this that makes use of materials he provides. Remember, for Kripke, possible states of the world are not alternate concrete universes, but abstract objects. They are maximally complete ways the real concrete universe could have been—maximally complete properties that the universe could have instantiated. Thinking of them in this way suggests an obvious generalization. Just as there are properties that certain objects could possibly have and other properties that they couldn't possibly have, so there are certain maximally complete properties that the universe could have had—possible states of the world—and other maximally complete properties that the universe could not have had—impossible states of the world.<sup>5</sup> If some of the properties that objects couldn't have had are, as our examples (1–4) indicate, properties that one can conceive them as having, then surely some maximally complete properties that the universe could not have had (some impossible states of the world) are properties one could conceive it as having. Given this, one could explain the informativeness of certain necessary truths as resulting from the fact that learning them allows one to rule out certain impossible, but nevertheless conceivable, states of the world. Moreover, one could explain the function played by empirical evidence in providing the justification needed for knowledge of the necessary propositions expressed by sentences (1–4); empirical evidence is required to rule out certain impossible, but nevertheless conceivable and epistemologically relevant, world-states with respect to which the propositions are false.<sup>6</sup> Thus, by expanding the range of epistemically conceivable states of the world to include some that are metaphysically impossible, one can modify the original conception of inquiry so as to accommodate Kripkean examples like (1–4). Whether or not the modification is sufficient to save the conception from further problems is a question to which we will return later.

<sup>5</sup> For an explanation and defense of this conception of world-states (in response to a different problem), see Nathan Salmon, "The Logic of What Might Have Been," *The Philosophical Review* 98, 1 (1989): 3–34.

<sup>6</sup> Here, I assume that names and indexicals (unlike certain definite descriptions such as *the stuff out of which this table, if it exists, is constituted*) rigidly designate the same thing with respect to all world-states, possible or not.

## Identity Statements and the Necessary A Posteriori

### *The Gap in Kripke's Argument*

I now turn to what may be the most well-known examples that Kripke gives of the necessary a posteriori—identity statements involving rigid designators.

5. Hesperus is Phosphorus.

6. Hesperus is Hesperus.

7.  $a = b$ .

If a sentence of the form (7) is true, then the term occupying the position of 'a' and the term occupying the position of 'b' refer to the same thing. If they are rigid designators, this means that they refer to the same thing with respect to all world-states in which that thing exists (and never refer to anything else). To keep things simple, let's ignore world-states in which the thing doesn't exist. Then we know that any true sentence of the form (7) in which the terms are rigid designators is a necessary truth; so (5) and (6) are necessary truths. (We take *is* in (5) and (6) to be the *is* of identity.) Of course, if we replace one of the names with a non-rigid designator, as in (8), then the example may be true without being necessary.

8. The planet seen at a certain place in the evening sky in certain seasons is Hesperus.

In lecture 2 of *Naming and Necessity*, Kripke considers an objection to this view which he attributes to Quine.

There's a dispute about this between Quine and Ruth Barcan Marcus. Marcus says that identities between names are necessary. If someone thinks that Cicero is Tully, and really uses 'Cicero' and 'Tully' as names, he is thereby committed to holding that his belief is a necessary truth. She uses the term 'mere tag'. Quine replies as follows, 'We may tag the planet Venus, some fine evening, with the proper name "Hesperus". We may tag the same planet again, some day before sunrise, with the proper name "Phosphorus". When we discover that we have tagged the same planet twice our discovery is empirical.'<sup>7</sup>

<sup>7</sup> *Naming and Necessity*, p. 100.

Quine concludes from this that the statement expressed by (5) must be contingent—presumably because it is not knowable a priori. Kripke's reply is, essentially, that this objection rests on the incorrect assumption that necessity and apriority come to the same thing. Once this is recognized, the objection collapses.

There is, however, a slightly different objection that could be read into Quine's comment about our "discovering that we have tagged the same planet twice." This phrase suggests that perhaps what he has in mind is (9).

9. 'Hesperus' and 'Phosphorus' refer (in our language) to the same thing.

The proposition expressed by (9) really is contingent, in addition to being knowable only a posteriori. However, the contingency of (9) is no argument against the necessity of (5), since the two sentences clearly express different propositions. Thus, Kripke's position remains intact.

Having considered these objections to his view that true identity statements involving names are necessary and a posteriori, Kripke devotes the last four pages of lecture 2 (the bottom of page 101 to the top of page 105) to explaining in detail why his view is correct. The section is too long to quote, so I will summarize it. The view Kripke presents 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 special names like *Hesperus* and *Phosphorus* the understanding of which may involve associating them with reference-fixing descriptions. In either case, Kripke argues, the evidence available to a competent user of the names is insufficient to determine that they are coreferential. He illustrates this point by noting that there is a possible state of the world in which speakers are in an evidentiary situation that is qualitatively identical with the one which we actual speakers are in, and yet, in the merely possible situation, the names are used to refer to different things. For example, there is a world-state with respect to 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 world-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 regarding their uses of

the names as we are. Yet their uses of the names refer to different things.

Kripke describes the case as follows:

The evidence I have before I know that Hesperus is Phosphorus is that I see a certain star or 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 in that sense we can say that it might have turned out either way.<sup>8</sup>

Kripke intends this example to show that the evidence available to us in the actual state of the world, as well as to agents in qualitatively similar world-states, 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 metaphysically 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. Kripke may have been thinking that for any proposition p, if it is not the case that p is true in all members of E, then p 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, and cannot, know apriori, simply on the basis of our mastery of the relevant terms or concepts. Let us grant this for the sake of argument. Well, one proposition that fails to be true in all members of E is the proposition that the names *Hesperus* and *Phosphorus* are coreferential; another closely related proposition is the one expressed by (10).

<sup>8</sup> *Ibid.*, pp. 103–4.



10. The sentence 'Hesperus is Phosphorus' expresses a truth in our language.

Thus, Kripke is in a position to conclude that the metalinguistic claim which says that the sentence 'Hesperus is Phosphorus' expresses a truth in our language is something that we cannot know apriori, but rather can come to be known only on the basis of empirical investigation.

So far so good. However, there is a problem. The lesson Kripke explicitly draws from this example is **not** that a certain metalinguistic claim is knowable by us only aposteriori, but rather that the claim that Hesperus is Phosphorus is knowable by us only aposteriori.

So two things are true: first, that we do not know *a priori* 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 the two planets in the sky, without the planets being the same.<sup>9</sup>

The problem is that Kripke's conclusion does not follow from his apparent premises. The proposition that Hesperus is Phosphorus is, as he insists, true in all metaphysically possible world-states. So it is true in all members of the class E of such world-states in which agents are in an epistemic situation qualitatively identical to ours. Since it is true with respect to those world-states, the principle that only propositions true in all members of E are knowable by us apriori does not rule out that we may know it apriori, even though it does rule out that we can know apriori that 'Hesperus is Phosphorus' expresses a truth in our language. Since the proposition expressed by (5) is **not** the same as the proposition expressed by (10), showing the latter to be knowable by us only aposteriori is not enough to establish that the former is knowable only in the same way.

The point I am making depends on sharply distinguishing (5), on the one hand, from (10), on the other. When explaining the **necessity** of (5), Kripke uses his example of possible world-states in which agents are in an epistemic situation qualitatively identical to ours to remind us that the contingency of (10) is irrelevant to the necessity of (5). According to Kripke, the agents in these world-states use the sentence 'Hesperus is Phosphorus' to express a **different** proposition from the proposition we actually use it to express. The fact that the proposition

<sup>9</sup> *Ibid.*, p. 104.

they use it to express is false with respect to their states of the world does not show that the proposition we actually use it to express is false at any world-state. What Kripke fails to point out is that the same reasoning can be applied to the epistemic status of the two examples. The proposition expressed by sentence (10) is knowable only *a posteriori*. But how does that bear on the question of whether proposition (5) is knowable *a priori*? The agents of Kripke's imagined world-states do not know the proposition **they** use the sentence 'Hesperus is Phosphorus' to express, for the simple reason that the proposition that they use it to express is **false** with respect to their world-states. But how does this show that the **different** proposition that **we** use the sentence to express isn't known by us, or that it isn't known by us independent of empirical investigation? Until we can answer this question, we have no way of viewing Kripke's discussion as supporting his conclusion that the claim that Hesperus is Phosphorus is not knowable *a priori* and so qualifies as an example of the necessary *a posteriori*.

#### *Kripke's Informal Strategy for Filling the Gap*

Although there is undoubtedly a gap in Kripke's argument, it is not an inexplicable one. Throughout the passage, he exploits a very familiar and highly intuitive connection between speakers' understanding and acceptance of sentences and our ability to use those sentences to report what they believe. For example, if you know that I fully understand and sincerely accept the sentence *Trenton is a town in central New Jersey*, then normally you will feel justified in reporting that I believe that Trenton is a town in central New Jersey. Similarly, if you know that I fully understand the sentence but do not accept it—either because I am uncertain of its truth or because I believe it to be false—then normally, and all other things being equal, you will feel justified in reporting that I **don't** believe that Trenton is a town in central New Jersey. The same is true of the sentence *Hesperus is Phosphorus*. In Kripke's example, before we made or learned of the astronomical discovery, we understood but did not accept the sentence; hence it is natural, and all but inevitable, to conclude that at that time we did not believe that Hesperus is Phosphorus. Moreover, as Kripke emphasizes with his evocation of agents in circumstances qualitatively identical to ours, we would not have been justified in accepting the sentence *Hesperus is Phosphorus* based on the evidence we had at that time. Because of this, it is natural to conclude that we wouldn't have been justified in believing

that Hesperus is Phosphorus based on the evidence we then had. But if that is so, then knowledge that Hesperus is Phosphorus must require empirical justification, in which case it must not be knowable a priori that Hesperus is Phosphorus—exactly as Kripke concludes.

This, I believe, is what Kripke had in mind. Although there is much to be said for this pattern of reasoning, there is a potential puzzle lurking within it. In order to bring out the puzzle, I will first use our modest theoretical apparatus of propositions as objects of belief and knowledge to formulate explicit and generalized premises to fill the gap in Kripke's argument. Next, I will scrutinize these new premises and show that either their truth is highly doubtful, or they are not strong enough to allow Kripke to derive his conclusion. Having isolated sources of potential difficulty, I will return to the informal reasoning just sketched, expose the puzzle that lies behind it, and evaluate Kripke's conclusion about this putative example of the necessary a posteriori.

### *A Formal and Explicit Reconstruction of Kripke's Argument*

The informal pattern of reasoning I have attributed to Kripke posits a close connection between our attitudes toward sentences and the beliefs those sentences are used to express. Using the terminology of propositions, we may express this thought as follows: Since sentences are vehicles for expressing propositions, the cognitive attitudes (belief, knowledge, etc.) that we bear to propositions expressed by sentences are mediated by the attitudes we bear to sentences that express them. Often, our believing a certain proposition goes hand in hand with understanding and accepting a sentence that expresses it. One initially plausible conception of the systematic connection between understanding and accepting a sentence, on the one hand, and believing the proposition it expresses, on the other, is stated by the following "strong disquotation" principle.<sup>10</sup>

#### STRONG DISQUOTATION

A sincere, reflective, rational individual *i* who understands a sentence *S* is disposed to accept *S*, and believe *S* to be true, iff *i* believes the proposition semantically expressed by *S*. Thus, if *S* is a sentence of English, then a reflective, rational

<sup>10</sup> To avoid complications, we will understand this principle as restricted to sentences that don't contain indexicals like 'I', 'now', and so on.

individual *i* who understands *S* satisfies the formula *x believes that S* iff *i* accepts *S*, and believes *S* to be true; if *S* is not a sentence of English, but is translatable into English as *P* (where *S* and *P* express the same proposition), then such an individual *i* satisfies *x believes that P* iff *i* accepts *S*, and believes *S* to be true.

Agents in an epistemic situation similar to ours before the astronomical discovery don't accept the sentence *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, according to the principle, 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 sentence 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*, and believing it to be true, on the basis of *e* iff *i*'s possession of *e* is sufficient to ensure that *i* would be justified in believing the proposition semantically expressed by *S*. So, if *S* is a sentence of English, then *i*'s possession of *e* is sufficient to ensure that *i* satisfies the formula *x would be justified in believing that S* iff *i* would be justified in accepting *S*, and believing *S* to be true, on the basis of *e*; if *S* is not an English sentence, but is translatable into English as *P* (where *S* and *P* express the same proposition), then *i*'s possession of *e* is sufficient to ensure that *i* satisfies *x would be justified in believing that P* iff *i* would be justified in accepting *S*, and believing it to be true, on the basis of *e*.

If these two principles are accepted, then Kripke's argument at the end of lecture 2 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 apriori, then it would be justifiable by the evidence available to us by virtue of our linguistic competence, plus our reasoning correctly about it.

(iv) So that belief is not justifiable apriori. Hence, it is not knowable apriori that Hesperus is Phosphorus.

This reconstruction of Kripke's argument has the virtue of being logically valid; the conclusion that it is not knowable apriori that Hesperus is Phosphorus is a logical consequence of Kripke's premises about the attitudes of agents toward the sentence *Hesperus is Phosphorus* plus the supplementary premise of strong disquotation and justification. Nevertheless, the argument is problematic—in part because there are reasons to suspect that the principles of strong disquotation and strong disquotation plus justification are untrue. One such reason is that these principles have the consequence that in order to believe a proposition, one must be disposed to accept every sentence one understands that semantically expresses it. Thus, the principles leave no room for the possibility that an individual might understand two sentences that semantically express the same proposition, without knowing that they do, and so might accept one of the sentences while not accepting the other. (In such a case the strong disquotation principle leads to the contradictory conclusion that the agent both believes and does not believe one and the same proposition.) Since there is reason to think that such possibilities are genuine, there is reason to reject the strong disquotation principles.<sup>11</sup>

One example of the problem with the strong disquotation principle is brought out by Kripke himself in his example of puzzling Pierre, discussed in his paper, "A Puzzle about Belief," published in 1979,

<sup>11</sup> For further discussion, see pp. 10–13 and chapter 3 of *Beyond Rigidity*. See also Nathan Salmon, "A Millian Heir Rejects the Wages of Sinn," in C. A. Anderson and J. Owens, eds., *Propositional Attitudes: The Role of Content in Logic, Language, and Mind* (Stanford, CA: CSLI, 1990), 215–47, at pp. 220–22. Also, Stephen Rieber, "Understanding Synonyms without Knowing That They Are Synonymous," *Analysis* 52 (1992): 224–28.

nine years after he delivered the lectures that became *Naming and Necessity*.<sup>12</sup> Kripke's Pierre is a Frenchman who has grown up in Paris speaking French, who 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 immersion, and comes to live in a poor and 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* name 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 his monolingual English-speaking neighbors in London do, who surely count as competent speakers. Moreover, the sentences are translations of one another; they mean the same thing and, it would seem, express the same proposition. But now we have a problem. By the strong disquotational principle (right to left direction) we get the result that Pierre does **not** believe the proposition expressed by *London is pretty*, since he understands but does **not** accept *London is pretty*. By a corresponding application of the strong disquotational principle (left to right direction) we get the result that he **does** believe the proposition expressed by *Londres est jolie*. Since the standard translation from French to English tells us that these sentences express the same proposition, we get the conclusion that Pierre both does and does not believe one and the same proposition. Since this is a contradiction, we have a *reductio ad absurdum* of the conjunction of the principle of strong disquotation with the standard (meaning-preserving) translation of French to English. This provides reason to doubt the strong disquotational principles.

<sup>12</sup> Saul Kripke, "A Puzzle about Belief," in A. Margalit, ed., *Meaning and Use* (Dordrecht: Reidel, 1979); reprinted in N. Salmon and S. Soames, eds., *Propositions and Attitudes* (Oxford: Oxford University Press, 1988). Kripke's formulation of the strong disquotational principle in that article is more informal and less theoretically loaded than mine. Although these differences are not without consequence, and are well worth studying, I will not discuss them here. Suffice it to say that they do not affect the difficulties we are about to uncover.

One natural response to this problem is to replace the strong disquotational principles—which are biconditionals—by the following weak disquotational principles—which are merely conditionals.

#### WEAK DISQUOTATION

If a sincere, reflective, rational individual *i* who understands a sentence *S* is disposed to accept *S*, and believe it to be true, then *i* believes the proposition semantically expressed by *S*. If *S* is a sentence of English, *i* thereby satisfies the formula *x believes that S*; if *S* is not a sentence of English, but is translatable into English as *P* (where *S* and *P* express the same proposition), then *i* satisfies the formula *x believes that P*.

#### WEAK DISQUOTATION AND JUSTIFICATION

If a sincere, reflective, rational individual *i* who understands *S* and is in possession of evidence *e* would be justified in accepting *S*, and believing *S* to be true, on the basis of *e*, then *i*'s possession of *e* is sufficient to ensure that *i* would be justified in believing the proposition semantically expressed by *S*. So, if *S* is a sentence of English, then *i*'s possession of *e* is sufficient to ensure that *i* satisfies the formula *x would be justified in believing that S*, if *i* would be justified in accepting *S*, and believing *S* to be true, on the basis of *e*; if *S* is not an English sentence, but is translatable into English as *P* (where *S* and *P* express the same proposition), then *i*'s possession of *e* is sufficient to ensure that *i* satisfies *x would be justified in believing that P*, if *i* would be justified in accepting *S*, and believing *S* to be true, on the basis of *e*.

When the principle of weak disquotation is substituted for that of strong disquotation in the story about Pierre, we no longer get the contradictory result that Pierre believes and does not believe the same thing. Instead, we get the weaker, and potentially defensible, result that he has contradictory beliefs; he believes that London is pretty by virtue of understanding and accepting the French sentence *Londres est jolie*, and he believes that London is not pretty (by virtue of understanding and accepting the English sentence *London is not pretty*).<sup>13</sup>

<sup>13</sup> Although Kripke himself draws no firm conclusion about the correctness or incorrectness of the principle of weak disquotation based on this case, he does find it puzzling and problematic that the principle leads to the result that Pierre—who may be fully rational and logical—has

However, even if this result is acceptable, and the principle of weak disquotation and justification is substituted for that of strong disquotation and justification in our reconstructed version of Kripke's argument, the argument still does not go through. The problem now is that the conclusion no longer follows from the premises. In this argumentative scenario, all that follows from our not being justified in accepting the sentence (5), *Hesperus is Phosphorus*, solely on the basis of our understanding it, is that **if** the proposition it expresses can be known, or justifiably believed, apriori, then such knowledge or belief must arise from something other than understanding and accepting that very sentence. This leaves open the possibility that there might be some sentence other than (5) which both expresses the proposition that Hesperus is Phosphorus, and is such that one can be justified in accepting it, and believing it to be true, solely on the basis of one's understanding of it. Since nothing in our new reconstruction of Kripke's argument rules out this possibility, the version of the argument employing the weak disquotational principles does not entail his conclusion. Although the conclusion is entailed by the version of the argument employing the strong disquotational principles, there is little comfort in this, since, as we have seen, they appear to be false.

### *The Lesson to be Learned: The Nontransparency of Meaning*

Although the two formal reconstructions of Kripke's arguments fail in different ways, the source of the failure is the same in both cases—the nontransparency of meaning. Two sentences may mean the same thing, and hence semantically express the same proposition, even though a competent speaker who understands both sentences does not realize this, and so accepts one, and believes it to be true, while refusing to accept the other, and either believing it to be false, or suspending judgment on it. This was Pierre's situation with the sentences *Londres est jolie* and *London is pretty*. Since the principle of strong disquotation is incompatible with this seemingly obvious truth, it should, I believe, be rejected—in which case the first reconstruction of Kripke's argument must be judged unsound.

---

contradictory beliefs. For a defense of both weak disquotation and this result, see Nathan Salmon, *Frege's Puzzle* (Cambridge, MA: MIT Press, 1986), and my "Direct Reference, Propositional Attitudes, and Semantic Content," originally published in *Philosophical Topics* 15 (1987): 47–87; reprinted in Salmon and Soames, *Propositions and Attitudes*.



The second reconstruction of the argument—relying on weak disquotation—fails because it does not rule out the possibility that there might be some sentence *S* other than *Hesperus is Phosphorus* which semantically expresses the same proposition as it does, and which is such that understanding *S* is sufficient for accepting it and knowing it to be true. How, one might wonder, could there be such a sentence? Well, think again about names. If, as Kripke contends, names do not have descriptive meanings or senses, what do they mean? One natural thought is that the meaning of a name is its referent. If this thought is correct, then coreferential names like *Hesperus* and *Phosphorus* mean the same thing. But surely if two names mean the same thing, then substitution of one for another in a simple sentence like (5) should preserve the meaning of the sentence. By this reasoning, we may reach the conclusion that the sentences *Hesperus is Phosphorus* and *Hesperus is Hesperus* mean the same thing, and hence semantically express the same proposition. Surely, it is reasonable to think that understanding *Hesperus is Hesperus* is sufficient for accepting it, and coming to know that it expresses a truth. Thus, it is natural to conclude that it is knowable a priori that Hesperus is Hesperus. But then, if the proposition that Hesperus is Phosphorus is nothing more than the proposition that Hesperus is Hesperus, it must be knowable a priori, contrary to Kripke.

Although this line of reasoning may initially seem far-fetched, and although it certainly goes beyond Kripke, there is nothing in *Naming and Necessity* to refute it. Moreover, once it has been recognized that meaning is not always transparent, one cannot object to the claim that *Hesperus is Hesperus* and *Hesperus is Phosphorus* mean the same thing—and hence semantically express the same proposition—simply by observing that it is possible for someone to understand both sentences, and yet not know that they mean the same thing. This surely is possible, but since meaning isn't transparent, this possibility doesn't establish that the sentences mean different things. Without establishing this, there is no way to use our disquotational principles to get Kripke's desired result.

Of course, Kripke didn't state any principles of strong or weak disquotation in *Naming and Necessity*, nor did he invoke our modest theoretical apparatus of propositions as objects of belief and knowledge. Hence, one might wonder about the relevance of our findings for his informal argument at the end of lecture 2. As we shall see, there is cause for concern.

His informal argument that it is not knowable apriori that Hesperus is Phosphorus is roughly the following:

(i) For someone who understands the sentence *Hesperus is Phosphorus*, sincerely accepting it and believing it to be true goes hand in hand with believing that Hesperus is Phosphorus—i.e., one who understands *Hesperus is Phosphorus* accepts it and believes it to be true iff one believes that Hesperus is Phosphorus.

(ii) Similarly, one who understands the sentence *Hesperus is Phosphorus* would be justified in accepting it and believing it to be true iff one would be justified in believing that Hesperus is Phosphorus.

(iii) In order to be justified in accepting *Hesperus is Phosphorus* and believing it to be true, it is not sufficient for one simply to understand it; in addition, one needs empirical evidence that the two names refer to the same thing.

(iv) Therefore, understanding the sentence *Hesperus is Phosphorus* is not sufficient for one to be justified in believing that Hesperus is Phosphorus; in order for one who understands the sentence to be justified in believing that Hesperus is Phosphorus, one must have empirical evidence that the two names refer to the same thing.

(v) Therefore, the statement that Hesperus is Phosphorus is not knowable apriori.

Principles (i) and (ii) are informal Kripkean counterparts of specific instances of the principles of strong disquotation and strong disquotation and justification. Thus, in order to accept Kripke's informal argument, one must have good reason to believe that these principles are true.<sup>14</sup>

<sup>14</sup> I do think that this much strong disquotation was implicit in Kripke's argument in *Naming and Necessity*. With this in mind, consider the following passage from footnote 44 of "A Puzzle About Belief": "some earlier formulations expressed disquotationally such as 'It was once unknown that Hesperus is Phosphorus' are questionable in light of the present paper (but see the previous note for this case)." The point here is that although it is unquestionable that it was once unknown that *Hesperus is Phosphorus* expressed a truth, we now see that we should not jump to the conclusion that it was once unknown that Hesperus is Phosphorus. This is questionable because it may just be that to know that Hesperus is Hesperus is to know that Hesperus is Phosphorus. But if that is so, then the argument at the end of lecture 2 of *Naming and Necessity* is undermined. I believe Kripke did not see this possibility at the time of the lectures. Nevertheless, footnote 44 continues: "I was aware of this question at the time *Naming and Necessity* was written, but I did not wish to muddy the waters further than necessary at that time. I regarded the distinction between epistemic and metaphysical necessity as valid in any case and

Let's just concentrate on (i), which is a particular instance of the following more general principle.

- (i\*) Let  $S$  be an non-indexical sentence of English and let  $i$  be a competent speaker of English who understands  $S$ . Then, the attitude ascription  $x$  *believes that*  $S$  is true of  $i$  iff  $i$  accepts  $S$ .

Here is an example scenario that creates problems for (i\*):

A student at the local college, Martin Martin, is both the quarterback of the football team and the best math student in school. His math teacher, Professor McX, has graded his work and consulted other professors about him. On that basis, she forms the opinion that Martin is a brilliant mathematician, which she expresses by saying *Martin is a brilliant mathematician*. Since her class is large, she doesn't know him by sight, so when she decides to attend one of the football games on Saturday, and sees him turn in an astounding performance on the field, she does not realize that the talented quarterback she is watching is her student, Martin. Sitting in the stands with her friend Harriet, she makes numerous comments about Martin—pointing him out and saying things like *He is a wonderful athlete*. Later Harriet, who knows Martin's girlfriend, reports that Professor McX thinks that Martin is a wonderful athlete. Surely, Harriet's report is true; the professor does believe this. Yet, if one were to ask her: *Do you think that Martin is a wonderful athlete?* she would understand the question, but she would not assent. Although she understands the sentence *Martin is a wonderful athlete*, she has no basis for accepting it, since she doesn't realize that *Martin* is the name of the quarterback who impressed her so much. So, although she believes that Martin is a wonderful athlete, and although she understands the sentence *Martin is a wonderful athlete*, she does not accept it or believe it to be true.

---

adequate for the distinctions I wished to make." Three points: (i) although the distinction between different types of necessity (and possibility) is relevant for explaining many instances of the necessary a posteriori, it won't help here if knowing that Hesperus is Hesperus turns out to be the same as knowing that Hesperus is Phosphorus; (ii) in Kripke's argument at the end of lecture 2, he does not invoke the distinction between different kinds of necessity (or possibility), and (iii) although he may have been generally aware of the difficulties posed by Pierre-type examples for strong disquotational principles at the time of *Naming and Necessity*, it seems evident that he did not focus on their implications for his argument at the end of lecture 2, probably because at that time he did not take seriously the idea that *Hesperus is Hesperus* and *Hesperus is Phosphorus* might mean the same thing.

This scenario casts doubt on (i\*). But if (i\*) is in doubt, then there may be reason to doubt (i) and (ii) as well. We need not here try to confirm or resolve those doubts, or to decide precisely what to say about potentially problematic cases like the one just sketched. It is enough to have shown that the premises needed, and tacitly used, in Kripke's argument are insecure, and cannot, without further investigation, be taken for granted in establishing his conclusion that it is not knowable apriori that Hesperus is Phosphorus.<sup>15</sup>

As we have seen, this is true whether the premises are formally stated using the theoretical apparatus of propositions as semantic contents of sentences and objects of propositional attitudes, or whether they are more informally and less abstractly formulated as (i) and (ii). Hence, we have little choice but to conclude that, despite the fact that there are other quite legitimate examples of the necessary aposteriori given in *Naming and Necessity*, Kripke's discussion at the end of lecture 2 does **not** establish that the statements made by identity sentences involving coreferential names are among them.<sup>16</sup>

### *Final Assessment: A Dilemma*

Despite this negative result, I suspect that many may be prepared to 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, and that this shows that sentences which differ only in the substitution of coreferential proper names may mean

<sup>15</sup> My own view about this issue is that the sentence *Is it knowable apriori that Hesperus is Phosphorus?* may be used to ask different questions. If one uses it to ask whether the proposition semantically expressed by *Hesperus is Phosphorus* is knowable apriori, then the answer is *yes*. (For more, see below.) However, if one uses it to ask a question that could more explicitly be paraphrased as *Is it knowable apriori that the planet, Hesperus, seen in the evening, is the planet, Phosphorus, seen in the morning?* then the answer is: that proposition is not knowable apriori, but it is also not necessary. In other work, I develop the idea that sentences are often used to assert more than their semantic contents, and sometimes these contents are not asserted at all. See *Beyond Rigidity*, and also "Naming and Asserting," in A. Szabo, ed., *Semantics and Pragmatics* (Oxford University Press, forthcoming). Following in the footsteps of my Ph.D. student Michael McGlone, I have come to believe that if this general idea is correct, then much of what is confusing about Kripke's puzzle about belief is due to the difficulty of keeping track of which propositions are being questioned and asserted in which contexts. For a recent note on this, see my "Saul Kripke, the Necessary Aposteriori, and the Two-Dimensionalist Heresy."

<sup>16</sup> We will return in chapter 17 to Kripke's thoughts about the example presented at the end of lecture 2 when we take up his problematic response to a general objection to the necessary aposteriori that he discusses in lecture 3.

different things and semantically express different propositions. And if the sentences  $a = a$  and  $a = b$  semantically 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 is too. So, even if we don't have an argument for the truth of *The proposition that  $a = b$  is knowable only a posteriori*, accepting the alleged datum would allow one to block the most obvious line of argument for the falsity of that claim.

Fair enough. However, there is still a difficulty to be faced. We need some positive theory of the contributions made by proper names to the propositions semantically expressed by sentences containing them. If the alleged datum is to be accepted, then this account must make clear precisely in what respect the propositions expressed by sentences containing different but coreferential proper names can, and in certain cases do, differ. What makes this task daunting is that the old solution to this problem—namely, the view that names have descriptive semantic contents—seems to have been thoroughly discredited by Kripke's arguments. If this is right, if the idea that names have descriptive semantic contents really has been discredited, then, given the alleged datum, one cannot identify the semantic contents of names either with their referents, or with descriptive semantic information that may vary from one coreferential name to another. In what other way do coreferential names differ? Often, they have different spellings, pronunciations, or syntactic structures, 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 whenever speakers use words with different spellings, pronunciations, or syntactic structures, they must, thereby, differ in the propositions they assert or believe.

Consequently, one who accepts Kripke's anti-descriptivist arguments is left with a dilemma. There are two main options. According to the first option, one may accept the alleged datum that, typically, when sentences differ only in the substitution of one coreferential proper name for another, it is possible to assert, believe, or know the proposition semantically expressed by one of the sentences without asserting, believing, or knowing the proposition semantically expressed by the other. If one does take this option, then one must give some positive account of propositions and propositional attitudes that explains how this is possible. That doesn't seem easy.

According to the second option, one may reject the alleged datum

and identify the semantic contents of names with their referents. If one does this, then one will be led to maintain that sentences which differ only in the substitution of coreferential names semantically express the same propositions, and, given this, one will find it very natural to conclude that attitude ascriptions involving such sentences are semantically equivalent. On this view, the proposition expressed by *Hesperus is Phosphorus* is identical with the proposition expressed by *Hesperus is Hesperus*. But then if (as many assume) the latter is knowable a priori, the former is too—in which case the claim that (5) is an example of the necessary a posteriori is false. This is the view defended by Nathan Salmon in *Frege's Puzzle*.<sup>17</sup> It is also the alternative I favor. However, it too faces a serious difficulty. The difficulty is to explain how, if the view is correct, speakers succeed in using sentences P and Q that differ only in the substitution of one coreferential proper name for another to assert and convey different information, and to express different beliefs—which they clearly do. In addition, it must be explained why speakers often do not regard an attitude ascription *n asserts/believes/knows that P* to be truth-conditionally equivalent to the corresponding attitude ascription *n asserts/believes/knows that Q*—which they often do not.

So the dilemma is either (i) to take it for granted that the sentences *Hesperus is Hesperus* and *Hesperus is Phosphorus* semantically express different propositions, and try to explain what this difference consists in, or (ii) to treat the two sentences as semantically expressing the same proposition, and try to explain how speakers use them to assert and convey different information, and to express and report different beliefs. The resolution of this dilemma is one of the most important unfinished legacies left to us by *Naming and Necessity*.<sup>18</sup>

### A Further Note on Inquiry

At the beginning of the chapter, we considered a conception of inquiry according to which the point of rational belief formation is to locate the actual state of the world by eliminating merely possible

<sup>17</sup> Salmon, *Frege's Puzzle*.

<sup>18</sup> In chapters 3 and 8 of *Beyond Rigidity*, I opt for (ii) and try to provide the basis for the needed explanations; these ideas are revised and extended in “Naming and Asserting.” See

world-states—i.e., maximally complete properties that the universe could have instantiated—incompatible with propositions one has come to know or believe on the basis of evidence. A revision of this conception was suggested in order to accommodate necessary *a posteriori* propositions expressed by sentences like (1–4). According to the revised conception, conceivable world-states—maximally complete properties that one can conceive the universe as instantiating (whether or not they could in fact be instantiated)—replace genuinely possible world-states in the model of inquiry. On this conception, the point of rational belief formation is to locate the actual state of the world by eliminating merely conceivable world-states incompatible with propositions one has come to believe on the basis of evidence. It is important to notice that our discussion in the previous section calls even this revised conception of inquiry into question. Suppose that it is possible to rationally believe a proposition *p*—like the proposition that London is pretty—by virtue of understanding and accepting a sentence *P* that expresses it, while also rationally believing the negation of *p*, by virtue of understanding and accepting the negation,  $\sim Q$ , of a different sentence—without being in a position, or having the evidence necessary, to recognize the inconsistency. In this sort of case, it would seem that the fact that the agent believes *p* does not eliminate all conceivable world-states with respect to which the negation of *p* holds. On the contrary, when the agent considers such states, as they are presented by the sentence  $\sim Q$ , they remain for him genuine epistemic possibilities. But if believing *p* doesn't eliminate the epistemic possibility that the negation of *p* may be true, then the notion of eliminating conceivable states **incompatible** with what one believes becomes problematic, and the model of inquiry based upon it breaks down. Thus, if inconsistent beliefs of the sort discussed here are genuinely possible, then a conception of inquiry substantially different from those we have considered may be needed.<sup>19</sup>

Salmon, *Frege's Puzzle*, for a different defense of (ii). For defenses of (i), see Mark Richard, *Propositional Attitudes* (Cambridge: Cambridge University Press, 1990), and Larson and Ludlow, "Interpreted Logical Forms."

<sup>19</sup> This point is related to arguments designed to show that propositions cannot be identified with sets of circumstances (metaphysically possible world-states, epistemically conceivable world-states, logically possible world-states, situations, etc.) with respect to which they are true. For such arguments, see my "Direct Reference, Propositional Attitudes, and Semantic Content."

# CHAPTER 16

## THE CONTINGENT APRIORI

### CHAPTER OUTLINE

#### 1. *Names Introduced by Reference-Fixing Descriptions*

The way Kripke's examples of the contingent apriori depend on names the referents of which are semantically fixed by descriptions; the argument that sentences containing such names express singular propositions about their referents

#### 2. *Prerequisites of Reference-Fixing and the Contingent Apriori*

Scenario 1: Restrictions on Reference Fixing. Problematic consequences for the contingent apriori of requiring that speakers know of a certain object that it satisfies a description prior to using the description to fix reference

Scenario 2: Unrestricted Reference Fixing. Consequences for the contingent apriori of allowing unrestricted use of descriptions to introduce and semantically fix the referents of names

Problems with scenario 2. Why putative examples of the contingent apriori based on the unrestricted use of

descriptions to fix the referents of names are not compelling

Weak disquotation vs. unrestricted reference-fixing. The conflict between two principles needed to make Kripke's examples of the contingent apriori work; why reference-fixing by description must be restricted

Why no Kripke-style examples can be genuine instances of the contingent apriori

#### 3. *Are There Examples of the Contingent Apriori?*

The importance of Kripke's insight over and above the details of his execution; how to use the actuality operator to construct genuine instances of the contingent apriori

### **The Semantics of Names Introduced by Reference-Fixing Descriptions**

In this chapter, we turn to Kripke's discussion of the contingent apriori. As with our discussion of the necessary aposteriori, we adopt the expository framework given at the beginning of chapter 15, according to which propositions are the things expressed by sentences, the bearers of



(contingent or necessary) truth, as well as the objects of belief, knowledge, and assertion. In addition, we continue to assume that an attitude ascription *x asserts/believes/knows that S* reports a relation—assertion/belief/knowledge—between an agent and the proposition expressed by *S*. Kripke's central doctrine about the contingent apriori may then be understood to be the claim that there are a number of cases in which a single proposition is contingent (i.e., true with respect to the actual state of the world, and false with respect to some other possible world-states) while nevertheless being knowable apriori (without appeal to empirical evidence).

Kripke's examples of the contingent apriori in *Naming and Necessity* involve cases in which a name has its referent semantically fixed by a non-rigid description. In these cases, he contends, we introduce a name *n* by stipulating that it is to rigidly designate the unique thing that satisfies a certain descriptive condition *D*. If there is such a thing, and the stipulation is successful, then the referent of *n* with respect to an arbitrary world-state *w* will be the actual denotation of *the D*, even though *n* will typically not be synonymous with the description, and sentences containing *n* will typically differ in meaning from corresponding sentences containing *the D*. An example of this is provided by the sentences in (1) and (2).

- 1a. If *n* exists, then *n* is the *D*.
- b. If there is a unique thing which is *D*, then *n* is the *D*.
- 2a. If the *D* exists, then the *D* is the *D*.
- b. If there is a unique thing which is *D*, then the *D* is the *D*.

Although the sentences in (2) are necessary truths, (1a) is contingent whenever *the D* is non-rigid, and (1b) is contingent whenever the way in which *the D* is non-rigid is by denoting, with respect to certain merely possible world-states, objects other than the entity it actually denotes. However, despite being contingent, (1a) and (1b) are taken to be apriori, on the grounds that understanding them is supposed to be sufficient for concluding that they are true. Here, it is worth noting that although Kripke often talks about **sentences** of this sort being both contingent and apriori, he typically feels free to express his view using the language of **indirect discourse**, as well. For example, he is willing to say, *It is possible for one to know apriori that if there is a unique thing which is D, then n is the D, even though it is not a necessary truth that*

*if there is a unique thing which is D, then n is the D, or to say, The claim that if there is a unique thing that is D, then n is the D is knowable apriori even though it is a contingent rather than a necessary truth.*

The most prominent example of this sort in *Naming and Necessity* concerns the standard meter. Here, Kripke imagines using the description *the length of stick s in Paris at time t* to fix the referent of the term *one meter*. He suggests that one who does this can know, solely on the basis of this reference-fixing stipulation, **that stick s is one meter long at t (if s exists at t, and hence has a unique length at t)**. He takes this to be an example of the contingent apriori, since clearly, s (if it had been heated or cooled) could have been some length other than this length (imagine a visual demonstration of its actual length), which is the length that the term *one meter* rigidly designates. Here is what he says.

What then, is the *epistemological* status of the statement ‘Stick S is one meter long at  $t_0$ ’, for someone who has fixed the metric system by reference to stick S? It would seem that he knows it *a priori*. For if he used stick S to fix the reference of the term ‘one meter’, then as a result of this kind of ‘definition’ (which is not an abbreviative or synonymous definition), he knows automatically, without further investigation, **that S is one meter long**. On the other hand, even if S is used as the standard of a meter, the *metaphysical* status of ‘S is one meter long’ will be that of a contingent statement, provided that ‘one meter’ is regarded as a rigid designator: under appropriate stresses and strains, heatings or coolings, S would have had a length other than one meter even at  $t_0$ . (Such statements as ‘Water boils at 100°C at sea level’ can have a similar status.) So in this sense, there are contingent *a priori* truths.<sup>1</sup>

Although there is something quite plausible about what Kripke says, there is a puzzle about this example, and all others of the same type. In each of these cases, Kripke produces an example sentence S, and goes on to argue that there is, in effect, some one claim (statement or proposition) designated by the clause *that S* which is both contingently true and knowable apriori. But it is unclear precisely what claim (statement or proposition) this is. For example, consider the claim that if stick s exists at t, then its length at t is one meter. We are told that this claim

<sup>1</sup> *Naming and Necessity*, p. 56, my boldface emphasis.

is both contingent and knowable a priori. Since it is contingent, we know that it cannot be the claim that if stick *s* exists at *t*, its length at *t* is the length of stick *s* at *t*. This is an example of an elementary point that Kripke insists on: When the reference of *n* is semantically fixed by a description *the D*, then typically *n* is not synonymous with *the D*, and sentences that differ only in the substitution of one for the other typically mean different things, and express different propositions (make different claims).

So what is this claim that is supposed to be both contingent and a priori? What exactly is the proposition expressed by *If stick s exists at t, then its length at t is one meter*? Perhaps it is the claim that if stick *s* exists at *t*, then its length at *t* is **this length** (imagine a visual demonstration of the length that is the actual length of the stick). Perhaps the claim in question is a singular proposition that says of a certain length *l*—the length rigidly designated by the name *one meter*—that if stick *s* exists at *t*, then its length at *t* is *l*.<sup>2</sup> The question at issue is what names the referents of which have been fixed by descriptions contribute to the propositions expressed by sentences containing them. Though Kripke doesn't himself use this theoretical language, he has already implicitly told us that they don't contribute the senses of their reference-fixing descriptions. So perhaps what they contribute are just the objects they rigidly designate—a certain length *l*, in the case of *one meter*. Some support for this idea can be found in the paragraph immediately following the passage just cited.

In the case of names one might make this distinction too [between 'definitions' which fix a reference and those which give a synonym]. Suppose the reference of a name is given by a description or a cluster of descriptions. If the name *means the same* as that description or cluster of descriptions, it will not be a rigid designator. It will not necessarily designate the same object in all possible worlds, since other objects might have had the given properties in other possible worlds, unless (of course) we happened to use essential properties in our description. So suppose we say, 'Aristotle is the greatest man who studied with Plato'. If we used that as a *definition*, the name 'Aristotle' is to mean 'the greatest man who studied with Plato'.

<sup>2</sup> The relationship between these two conjectures is the following: the proposition expressed by *If stick s exists at t, then its length at t is this length*, said demonstrating the precise length *l* in question, either is **identical** with, or trivially **entails**, the singular proposition that says of *l* that that if stick *s* exists at *t*, then its length at *t* is *l*. For purposes of our discussion, it doesn't matter whether the relationship is identity of entailment. Hence, to simplify the presentation, I will often speak as if it were identity.

Then of course in some other possible world that man might not have studied with Plato and some other man would have been Aristotle. If, on the other hand, we merely use the description to *fix the referent* then that man will be the referent of 'Aristotle' in all possible worlds. The only use of the description will have been to pick out to which man we mean to refer. But then when we say counterfactually 'suppose Aristotle had never gone into philosophy at all', we need not mean 'suppose a man who studied with Plato, and taught Alexander the Great, and wrote this and that, and so on, had never gone into philosophy at all', which might seem like a contradiction. **We need only mean, 'suppose that *that man had never gone into philosophy at all*'.**<sup>3</sup>

Note, Kripke is quite explicit that if a description *the D* merely fixes the referent of a name *n*, then uses of *n is F* and *the D is F* may mean quite different things, formulate different claims, and so, it would seem, express different propositions. Moreover, he suggests that in these cases if *n* refers to some man, then someone who uses *n is F* means and says something like *that man is F*, where the term *that man* is used demonstratively to pick out the referent of *n*. This is interesting. Although he is not explicit about the significance of this point, his remarks invite a pair of inferences.

(i) Even when *n* is a name the referent of which is semantically fixed by a description, the content of one's use of *n is F* is a singular proposition that says, of the referent of *n*, that it has a certain property (the one expressed by *F*).

(ii) In order for one to know this proposition, and hence to satisfy the ascription *x knows that n is F*, one must know, of the particular object which is the referent of *n*, that it has a certain property (the one expressed by *F*).<sup>4</sup>

<sup>3</sup> Ibid., p. 57, my boldface emphasis.

<sup>4</sup> Here, and throughout, I use the locution *believe of an object that it is so and so*. To say that an agent *believes of o that it is F* is to say that the agent satisfies the formula *believes that x is F* relative to an assignment of *o* to '*x*'. This will be the case iff the agent believes the singular proposition that predicates the property expressed by *F* of *o*. Of course, in order to believe this proposition, one must think about *o* in a certain way, which may involve attributing various descriptive characteristics to *o*. Since these are not part of the singular proposition expressed by *x is F* relative to an assignment of *o* to '*x*', belief in that proposition is often accompanied by belief in other, descriptively enriched propositions. However, the truth of the claim *The agent believes of o that it is F* does not depend on which other descriptively enriched propositions the agent may also believe. For more on believing of—often called *de re belief*—see pp. 149–153 of my "Donnellan's Referential/Attributive Distinction," *Philosophical Studies* 73 (1994): 149–168.

These claims—(i) and (ii)—gain further plausibility from the observation that corresponding claims also seem to hold in the case of ordinary names, the referents of which are **not** semantically fixed by descriptions. For example, consider the name *Saul Kripke*, the referent of which is not simply stipulated to be the same as that of some reference-fixing description. Corresponding to (i), we observe that someone who assertively utters the sentence *Saul Kripke is F* says of a certain man, Saul Kripke, that he has a certain property (the one expressed by F). Corresponding to (ii), we notice that if Mary knows that Saul Kripke is a famous philosopher who was a native of Omaha, Nebraska, then there is a particular individual, Saul Kripke, such that she knows that he is a famous philosopher who was a native of Omaha, Nebraska—Mary knows of this particular man that a certain fact about him holds.

By contrast, principles corresponding to (i) and (ii) do not hold when descriptions are substituted for names. For example, suppose that Mary knows that there are spies, without knowing of any particular individual that he or she is a spy. Suppose further that she correctly believes that some one among the spies must be the shortest spy, again without having any idea who that person may be.<sup>5</sup> Then, contrary to (i), if Mary assertively utters the sentence *The shortest spy is a spy*, she has not said, of the individual Boris (who, unknown to her, is the shortest spy), that he is a spy, and, contrary to (ii), although Mary knows that the shortest spy is a spy, there is no particular person *p* such that Mary knows that *p* is a spy—i.e., there is no individual such that Mary knows, of that individual, that he or she is a spy. Thus, whereas (i) and (ii) hold for ordinary names, corresponding principles do not hold for many descriptions. In light of this, Kripke's implicit claim that (i) and (ii) also hold for special names the referents of which are semantically fixed by descriptions is a plausible and useful reminder that, for him, the propositions semantically expressed (and asserted) by (utterances of) sentences containing these special names are to be understood to be on a par with the singular propositions semantically expressed (and asserted) by (utterances of) sentences containing ordinary names.<sup>6</sup>

<sup>5</sup> This example is based on a distinction originally discussed by Quine in "Quantifiers and Propositional Attitudes," *Journal of Philosophy* 53 (1956), reprinted in Quine's *The Ways of Paradox*.

<sup>6</sup> On p. 41 of *Naming and Necessity* Kripke gives the same account of what is meant by an utterance of a sentence containing an ordinary name—*Nixon*—as he gives of what is meant by an utterance of a sentence containing a name the reference of which is fixed by a description. He says, "When you ask whether it is necessary or contingent that *Nixon* won the election, you are asking the intuitive question whether in some counterfactual situation, *this man* would in fact have lost the election." Note the demonstrative, and compare with the passage from p. 57 already cited.

In light of this, it is reasonable to think that, for Kripke, the proposition that if stick *s* exists at *t*, then its length at *t* is one meter is the very same proposition as the proposition that if stick *x* exists at *t*, then its length at *t* is this length (imagine a visual demonstration of the length which *one meter* rigidly designates). This proposition says of a certain length *l* that *l* is the length of stick *s* at *t*, if *s* exists at *t*. This, presumably, is the proposition that is supposed to be contingent, yet knowable a priori. More generally, Kripke's hypothesis may be taken to be the claim that all examples of the contingent a priori involving names introduced by reference-fixing descriptions are cases in which a single proposition is both contingent yet knowable a priori—where this proposition is a singular proposition, which says something of the particular individual that is the referent of the name.

### **Prerequisites of Reference-Fixing and Their Relevance for the Contingent A priori**

Although I believe this to be the correct reading of Kripke's text, it raises difficult interpretive and philosophical problems. In his discussion of the meter-stick example, which he takes to provide a genuine instance of the contingent a priori, he nowhere indicates whether, in fixing the referent of the term *one meter*, we have had independent acquaintance with stick *s*. This may encourage the reader to think that it doesn't matter for the philosophical points he is making whether or not we have had such contact. However, this is not so. In fact, a resolution of this issue is necessary, if we are to understand and evaluate the central points in Kripke's discussion of the contingent a priori.<sup>7</sup>

#### *Scenario 1: Restrictions on Reference-Fixing*

First, consider the scenario in which we have seen stick *s* at *t*, and formed an idea of its length, prior to introducing the term *one meter* with the description *the length of stick s at t*. In this scenario, we may be taken to have a perceptually justified true belief of a certain length, that *l* is the length of stick *s* at *t*. Consequently, when we introduce the term *one meter* as a rigid designator of that length, our understanding

<sup>7</sup> This issue is raised, and its significance insightfully discussed, by Nathan Salmon in "How to Measure the Standard Metre," *Proceedings of the Aristotelian Society* 88 (1987/88): 193–217.

of the sentence *the length of stick s at t is one meter (if s exists at t)* guarantees that we know both that it expresses a truth, and precisely which truth it expresses—we know that it expresses the true proposition that *l* is the length of stick *s* at *t* (if *s* exists at *t*). The only remaining question about this scenario is whether this result is sufficient for our knowledge of this proposition to qualify as apriori. It is not clear that it is. After all, in this scenario, we have a perceptually justified belief in this proposition even prior to the introduction of *one meter*. That belief is certainly not apriori; and it is hard to see why merely performing a linguistic ceremony should transform it into one. Hence it is hard to see how, in this scenario, we have a genuine example of the contingent apriori. We will return to this point later. Before we do, however, we need to examine a different way of understanding Kripke's example.

### *Scenario 2: Unrestricted Reference-Fixing*

In this scenario, we are portrayed as attempting to introduce the term *one meter* as a rigid designator of whatever satisfies the description *the length of stick s at t* in a situation in which we have never seen stick *s*, and have no idea how long it is. Suppose that it is possible to successfully introduce the term in this sort of case. Then the situation is one in which the following facts obtain.

- F1. There is a proposition *p* which truly says of a certain length—namely, this length (imagine a visual demonstration of the length *l* which is the actual length of the stick at *t*)—that it is the length of stick *s* at *t*, if stick *s* exists at *t*.
- F2. Prior to introducing the term *one meter*, we do not believe *p*, since we have no idea how long *s* is at *t*.
- F3. We introduce the term *one meter* by stipulating that it is to rigidly designate whatever length is the actual length of *s* at *t*.
- F4. As a result of this stipulation, the sentence *The length of stick s at t is one meter, if s exists at t* comes to express *p*.
- F5. We understand the sentence in the following sense: we understand the process of introducing a name by a description; we understand how the referent of the name is thereby determined; and we understand how the proposition expressed by the sentence is determined, given the referent of the name.

Understanding all this is sufficient for us to know that the sentence expresses a truth—i.e., that its meaning, which may be taken to be a function from contexts of use of the sentence to propositions expressed, assigns a true proposition as value to our context as argument.

At this point, two questions arise.

- Q1. How, if at all, does the knowledge mentioned in F5 result in our knowing  $p$ , and hence in our knowing that the length of stick  $s$  at  $t$  is one meter, if  $s$  exists at  $t$ ?
- Q2. If the knowledge mentioned in F5 does result in our knowing  $p$ , is this knowledge properly characterized as apriori knowledge?

Since *Naming and Necessity* is not explicitly committed to F1–F5, it does not provide explicit answers to Q1 and Q2. There is simply nothing in Kripke's discussion of the meter-stick example that definitively settles the matter of whether one can introduce a name by a reference-fixing description, without acquaintance with its denotation and independent evidence that the description applies to it. However, other parts of his discussion have encouraged some readers to think that one can use a description to introduce and semantically fix the referent of a name, even when this condition is not satisfied. One of these passages is the following, in which Kripke gives the example of the "name" *Jack the Ripper* as an instance of a real-life case in which the referent of a name may be semantically fixed by a description.

What picture of meaning do these Theses ((1)–(5)) give you? [The theses are meant to characterize cases in which the referents of names are semantically fixed by descriptions.] The picture is this. I want to name an object. I think of some way of describing it uniquely and then I go through, so to speak, a sort of mental ceremony: By 'Cicero' I shall mean the man who denounced Catiline; and that's what the reference of 'Cicero' will be. I will use 'Cicero' to designate rigidly the man who (in fact) denounced Catiline, so I can speak of possible worlds in which he did not. But still my intentions are given by first, giving some condition which uniquely determines an object, then using a certain word as a name for the object determined by this condition. **Now there may be some cases in which we actually do this.** Maybe, if you want to stretch



and call it description, when you say: I shall call that heavenly body over there ‘Hesperus’. **That is really a case where the theses not only are true but really even give a correct picture of how reference is determined. Another case, if you want to call this a name, might be when the police in London use the name ‘Jack’ or ‘Jack the Ripper’ to refer to the man, whoever he is, who committed all these murders, or most of them. Then they are giving the reference of the name by a description.**<sup>8</sup>

Here Kripke explains what it is to use a description to fix the referent of a name, and indicates that (although most names, like ‘Cicero’, do not have their referents fixed in this way), ‘Hesperus’ and ‘Jack the Ripper’ may really fit the picture. Since the situation he imagines in the case of ‘Jack the Ripper’ is one in which the police had little or no direct contact with the man denoted by the description allegedly used to fix reference, and since they seem not to have had independent evidence to believe of any specific man that the description applied to him, this passage might be read as suggesting that, according to Kripke, one can use a description to semantically fix the referent of a name, even in cases in which one is not acquainted with the denotation of the description, and has no independent evidence that the description applies to that individual.

Next, we apply this idea to the scenario in which we use *the length of stick s at t* to fix the referent of *one meter*, even though we have never seen *s* and have no idea how long *s* is. This brings us back to F1–F5, and the two questions raised about those alleged facts. F5 tells us that we understand the sentence *The length of stick s at t is one meter, if s exists at t*, and that this understanding is sufficient to conclude that it expresses a truth (in our language). The next question is, how, if at all, this ensures that we know the truth it expresses. Here I hypothesize implicit reliance (by some readers and perhaps even Kripke himself) on some version of the principle of weak disquotation cited in the discussion in chapter 15 of the necessary aposteriori, and repeated here in abbreviated form.<sup>9</sup>

#### WEAK DISQUOTATION

If a sincere, reflective, rational individual *i* who understands a sentence *S* is disposed to accept *S*, and believe *S* to be true,

<sup>8</sup> *Naming and Necessity*, pp. 79–80, my emphasis and additions in brackets.

<sup>9</sup> A different version of the principle might speak of accepting the meaning of *S* and believing that meaning to assign a true proposition to one’s present context. Our discussion is intended to be neutral as to variations such as this.

then *i* believes the proposition semantically expressed by *S*, and thereby satisfies the formula *x believes that S*.

The idea is that the knowledge mentioned in F5 leads us both to understand and to accept the sentence *The length of stick s at t is one meter, if s exists at t*, and believe it to be true; weak disquotation then ensures that we believe the contingent proposition *p* that it expresses. How, then, is it determined that this knowledge qualifies as apriori? Here we appeal to the principle of weak disquotation and justification cited in chapter 15, and repeated here in abbreviated form.

#### WEAK DISQUOTATION AND JUSTIFICATION

If a sincere, reflective, rational individual *i* who understands *S* and is in possession of evidence *e* would be justified in accepting *S*, and believing it to be true, on the basis of *e*, then *i*'s possession of *e* is sufficient to ensure that *i* would be justified in believing the proposition semantically expressed by *S*, and hence it is sufficient to ensure that *i* satisfies the formula *x would be justified in believing that S*.

The knowledge mentioned in F5 not only leads us to accept the sentence *The length of stick s at t is one meter, if s exists at t*, and believe it to be true, it also **justifies** this acceptance and belief. Since this knowledge is simply knowledge of meaning, the principal of weak disquotation and justification guarantees that our knowledge of semantic facts about the sentence justifies us in believing the proposition *p* that it expresses, and hence in believing that the length of stick *s* at *t* is one meter, if *s* exists at *t*. Given that *p* is true and that we are justified in believing it, one may conclude that we know *p*—we know that the length of stick *s* at *t* is one meter, if *s* exists at *t*—solely by virtue of our understanding the sentence that expresses this knowledge, and knowing semantic facts about it.<sup>10</sup>

What does this have to do with the apriori? The idea, I suppose, is something like the following:

#### WEAK LINGUISTICISM ABOUT THE APRIORI

If one knows a proposition *p* solely by virtue of understanding a sentence that expresses *p*, and knowing semantic facts about it, then one knows *p* apriori.

<sup>10</sup> For our purposes, I here assimilate knowledge to justified true belief. Although such an analysis fails in certain special cases, none of these are in the offing here, and we can ignore unnecessary complications.

Given all this, one can characterize the proposition that the length of stick *s* at *t* is one meter, if stick *s* exists at *t*, as a legitimate example of the contingent apriori—where one’s apriori knowledge of this proposition is seen as arising from the use of the description *the length of stick s at t* to fix the referent of *one meter*, in a case in which one has never seen *s*, and one has no idea, prior to introducing the term, how long *s* is.

### *Problems with Scenario 2*

#### PROBLEM 1

Although the story we have told yields a putative example of the contingent apriori that fits much of what Kripke says in *Naming and Necessity*, in my opinion, the story is not believable, and in the end, I will argue, even Kripke doesn’t believe it. The first problem is that the reasoning employed in scenario 2 mischaracterizes the apriori, and eviscerates the distinction between propositions knowable apriori and propositions knowable only aposteriori. Part of the problem lies in the principle of weak linguisticism about the apriori. If this principle is not taken to be part of a new stipulative definition of what we are to mean by *apriori*, but rather is understood as a hypothesis about the traditional notion of knowledge the justification of which does not rest on empirical evidence, then it is extremely doubtful that it is true.

This point is illustrated by the examples in (3) and (4).

- 3a. ‘Rabbits are animals’ is a true sentence of English iff rabbits are animals.
- b. The word ‘rabbit’ refers in English to an object iff that object is a rabbit.
- 4a. ‘Rabbits are animals’ is a true sentence of my language iff rabbits are animals.
- b. The word ‘rabbit’ refers in my language to an object iff that object is a rabbit.

The propositions expressed by these sentences are standardly not regarded as capable of being known apriori. First consider those in (3). Anyone who understands their meanings, and knows that they are sentences of English, has all the information needed to determine their truth. However, the propositions they express cannot be known independent of empirical evidence. If they could be so known, then

non-English speakers should be able to know them simply by reflecting on and reasoning about them—which of course they cannot. What separates English from non-English speakers is that the former are in possession of empirical evidence about what English words and sentences mean. This evidence plays two roles in examples like (3a,b)—(i) it allows English speakers to understand the sentences, and (ii) it justifies their belief in the truth of the sentences, and the propositions they express. It is because of this second role played by empirical evidence that these propositions are not knowable apriori. The same holds for the propositions expressed by (4a,b). The only difference between this case and the one involving (3) is that whereas an English speaker might doubt whether the language that he speaks is “English” (and so be in doubt about (3a) and (3b)), I can hardly doubt that the language I speak is my language. Thus, in the case of (4a,b), all that is required in order for me to know that these sentences (and the propositions they express) are true is that I understand them. However, since the evidence that goes into this understanding also justifies my belief in the propositions they express, my knowledge of these propositions is not apriori in the traditional sense.

What has been said about (3) and (4) carries over to the account of the sentence *The length of stick s at t is one meter, if s exists at t* given in scenario 2. The reference-fixer’s knowledge that the sentence expresses a truth in his language is based on his understanding of it, which in turn is rooted in his knowledge of certain empirical facts, including facts about the use of the description *the length of stick s at t* to introduce and semantically fix the referent of the term *one meter*. Since his knowledge of the truth of this sentence is not apriori, and since, according to the weak disquotation principles, his justification for believing the proposition *p* that the sentence expresses is inherited from his justification for believing the sentence to be true, his knowledge of *p*—i.e., his knowledge that the length of stick *s* at *t* is one meter, if *s* exists at *t*—is not apriori either. In all these cases—(3), (4), and the scenario 2 treatment of the meter-stick example—the contingent propositions known to be true are (or are supposed to be) knowable solely on the basis of understanding sentences that express them. If they are knowable in this way, then they may properly be regarded as trivial in an interesting way. However, it is doubtful that they should be counted as apriori in the traditional sense. Thus, it is doubtful that weak linguisticism about the apriori is true.

The problem is compounded when this principle is combined, as it is in scenario 2, with weak disquotation, weak disquotation and justification, and the unrestricted ability to use descriptions to introduce and

semantically fix the referents of names. When this is done, the class of propositions characterized as *knowable apriori* becomes far too expansive to be theoretically interesting. Let  $o$  be any object whatsoever, and let  $\{P_1, \dots, P_n\}$  be any set of properties the conjunction of which uniquely applies to  $o$ . Given the unrestricted ability to use descriptions to introduce and semantically fix the referents of names, one could always form a description *the D* denoting the object uniquely designated by the conjunction of the  $P_i$ 's, and use it to rigidly fix the reference of a name  $n$ —which would then refer to  $o$ . Finally, let  $p$  be any proposition that says of  $o$  that, if it exists, then it has one or more of the  $P_i$ 's. Using  $n$ , we could formulate a sentence expressing  $p$  which would be guaranteed to express a truth, and would be known by linguistically competent speakers to do so. Under the assumptions of scenario 2, this is sufficient for  $p$  to be characterized as capable of being known apriori. In other words, if the assumptions of scenario 2 were correct, then virtually every true proposition that predicates one or more properties of any object would qualify as capable of being known apriori. Surely, this cannot be.

#### PROBLEM 2

I have already expressed doubt about the principle of weak linguisticism about the apriori. In addition, the difficulty just noted points to a further problem with the combination of weak disquotation and the unrestricted ability to use descriptions to introduce and semantically fix the referents of names. Thus, there is reason to think that one or the other of these two doctrines must be false. Recall the five alleged facts which, in scenario 2, arise from the unrestricted ability to use descriptions to semantically fix reference.

- F1. There is a proposition  $p$  which truly says of a certain length—namely, this length (imagine a visual demonstration of the length  $l$  which is the actual length of the stick at  $t$ )—that it is the length of stick  $s$  at  $t$ , if stick  $s$  exists at  $t$ .
- F2. Prior to introducing the term *one meter*, we do not believe  $p$ , since we have no idea how long  $s$  is at  $t$ .
- F3. We introduce the term *one meter* by stipulating that it is to rigidly designate whatever length is the actual length of  $s$  at  $t$ .
- F4. As a result of this stipulation, the sentence *The length of stick  $s$  at  $t$  is one meter, if  $s$  exists at  $t$*  comes to express  $p$ .

- F5. We understand the sentence in the following sense: we understand the process of introducing a name by a description; we understand how the referent of the name is thereby determined; and we understand how the proposition expressed by the sentence is determined, given the referent of the name. Understanding all this is sufficient for us to know (i) that the sentence expresses a truth, and (ii) that its meaning, which may be taken to be a function from contexts of use of the sentence to propositions expressed, assigns a true proposition as value to our context as argument.

Adding weak disquotation to this collection gives the result that we believe *p*, and hence believe that the length of stick *s* at *t* is one meter, if *s* exists at *t*. But notice how mysterious this is. Priori to going through a little verbal ceremony, we are ignorant of a certain entirely nonlinguistic, nontrivial, empirical fact—namely that stick *s*, if it exists at *t*, is this long (I visually demonstrate the length *l*, which is its actual length at *t*). We say a few formulaic words—*let ‘one meter’ rigidly designate the length of stick s at t, whatever it might be*—and, presto, we know the fact that we were previously ignorant of. Wait until students taking exams hear about this. When asked *Who did such and such?*, they may answer *N did such and such, where ‘N’ is a name I hereby stipulate has its reference rigidly fixed by the description, ‘the individual who did such and such’*. Such a performance would be absurd. True belief is not that easy to come by, which suggests either (i) that there are substantial restrictions on when descriptions can be used to introduce and semantically fix the referents of names, or (ii) that there are no such restrictions, but weak disquotation fails when applied to sentences containing such names.

Another indication that this is the case is provided by a persuasive argument given by Kripke himself, but, as far as I know, one that he never published.<sup>11</sup> Let *A* be any agent who has at least one false belief, which is expressed by the sentence *S*. Next, consider the description (5).<sup>12</sup>

5. The *x*: (if *S*, then *x* = Princeton University) & (if  $\sim$ *S*, then *x* = Saul Kripke’s left thumbnail).

<sup>11</sup> I am familiar with the argument from his seminars at Princeton University in the ’80s and ’90s, as well as from some of his public lectures elsewhere. I do not know whether he was aware of the argument at the time he gave the lectures in *Naming and Necessity*.

<sup>12</sup> *The x: . . . x . . .* is a singular definite description formed by attaching the description operator *the x* to a formula containing at least one free occurrence of ‘*x*’. It is paraphrased *the unique individual x which is such that. . .*

Since S is false, this description refers to Kripke's left thumbnail. However, since A believes that S is true, he will take it to refer to Princeton University, and hence will accept sentence (6a), and wrongly believe it to express a truth.

- 6a. The x: (if S, then x = Princeton University) & (if ~S, then x = Saul Kripke's left thumbnail) is an institution of higher learning.

So far, there is nothing paradoxical about this. Having one false belief naturally leads to having other closely related false beliefs, and this is just an instance of that.

A problem arises, however, if it is possible for A to use (5) to semantically fix the referent of a name 'PU', even though A has no independent evidence that its actual denotation, Saul Kripke's left thumbnail, satisfies the description. Given this latitude, which is assumed in scenario 2, we get the result that A understands, accepts, and believes true the sentence (6b), which expresses a singular proposition p which says of Kripke's thumbnail that it is an institution of higher learning.

- 6b. PU is an institution of higher learning.

Combining this result with weak disquotation gives the result that A believes p, and hence that (7a) is true.

- 7a. A believes that PU is an institution of higher learning.

Since PU is Saul Kripke's left thumbnail, and since (7a) tells us that A believes of it that it is an institution of higher learning, we get (7b).

- 7b. There is a certain object o such that o is Saul Kripke's left thumbnail, and A believes that o is an institution of higher learning.

Kripke rightly regards this result—which could be repeated for virtually any agent, object, and property—to be a *reductio ad absurdum* of the conjunction of weak disquotation with the unrestricted ability to use descriptions to introduce and semantically fix the referents' names. Thus, it seems, at least one of these principles must go.

This is a result that ought to have been expected all along. The principle of weak disquotation is plausible, and makes good sense, so long as there is a close relationship between understanding a sentence and knowing what proposition it expresses. When there is such a relationship, someone who understands the sentence may be thought of as entertaining the proposition it expresses, and someone accepting the

sentence is plausibly taken as believing the proposition. If, however, one opens a substantial gap between understanding a sentence *s* and knowing of a certain proposition *p*, that *s* expresses *p*, then the original rationale for weak disquotation is undermined. This is just what happens when one assumes both that sentences containing names express singular propositions that predicate various things of their referents, and that there are essentially no restrictions on the use of descriptions to semantically fix the referents of names.

Under these assumptions, such names constitute a kind of indexical, with their referents at a given context of utterance being determined to be whatever satisfies their reference-fixing descriptions with respect to the possible world-state of the context of utterance. On this picture, knowing the referent of such a name *n* at a context, and knowing of a proposition *p* that it is expressed by the sentence containing *n*, requires knowing which object satisfies the description associated with *n* at the possible world-state of the context. As a result, an agent who lacks this knowledge will no more know which proposition is expressed by his own words than would an agent at Yankee Stadium who assertively uttered *He is a Yankees fan*, with his eyes closed while pointing blindly at me in a crowd, not having any idea at whom he was pointing.

Note, the agent with his eyes closed understands the sentence he utters in the sense that he knows what it means. He may also believe that his use of it expresses a truth, since he may regard it as very likely that the person he is pointing at is a Yankee's fan. However, it would be wrong to suppose that he thereby acquires a (false) belief about *me*. Since he has no idea at whom he is pointing, he does not believe that I am a Yankees fan. By the same token, he does not believe a priori the contingent truth that if I exist, then I am the person he is pointing at—which he would express by uttering *If he exists, then he is the person I am pointing at* while pointing blindly at me. Thus, it would be a mistake to suppose that weak disquotation correctly applied to such a case. By the same token, if one thought that there were essentially no restrictions on the use of descriptions to semantically fix the referents of names, then it would be a mistake to suppose that weak disquotation applied to sentences containing such names.<sup>13</sup>

<sup>13</sup> Though the case seems clear to me, not everyone agrees. For example, Gil Harman supports what is essentially the opposite point of view in his "How to Use Propositions," *American Philosophical Quarterly* 14 (1977): 173–76. See also section XVII of David Kaplan, "Demonstratives," in Almog, Perry, and Wettstein, *Themes From Kaplan*, and Robin Jeshion, "Acquaintanceless De Re Belief," in M. O'Rourke et al. eds., *Truth and Meaning: Investigations in Philosophical Semantics* (New York: Seven Bridges Press 2001).



*Weak Disquotation vs. Unrestricted Reference-Fixing*

At this point, we are left with two possible alternatives. According to view 1, descriptions may freely be used to semantically fix the referents of names, whether or not one has independent knowledge of which objects satisfy the descriptions, or which propositions are expressed by sentences containing such names. However, if one lacks such knowledge, then accepting these sentences, and believing them to be true, will not be sufficient for one to believe the propositions they express. On this view, names the referents of which are fixed by descriptions are easy to come by, but sentences containing them are not of much use for expressing beliefs, because weak disquotation fails for such sentences. Thus, when *the D* semantically fixes the referent of *n*, one may know, solely on the basis of one's understanding of the language, that the sentences *If there is a unique D, then n is the D* and *If n exists, then n is the D* express truths; however, one cannot know the propositions expressed by these sentences on this basis, and in many cases one may fail to know, or believe, them at all. According to view 2, a description cannot semantically fix the referent of a name for a speaker unless the speaker independently believes of the object which is denoted by the description that the description applies to it. However, if this condition is satisfied, then a speaker who understands, accepts, and believes the sentence to be true will believe the proposition the sentence expresses. Thus, when *the D* is successfully used to fix the referent of *n*, a speaker may know, solely on the basis of his understanding the language, the propositions expressed by the sentences *If there is a unique D, then n is the D* and *If n exists, then n is the D*, in addition to knowing that these sentences express truths. However, in this case, the knowledge required to understand the language will include the independent, non-linguistic knowledge reported by *x knows of y that it is D* relative to an assignment of the speaker to 'x', and the object denoted by *the D* to 'y'.

Faced with this choice, we have reason to prefer view 2. As I see it, view 1 simply creates too large a gap between the proposition expressed by a sentence in a context, and the information with which competent speakers are presented when they understand the sentence, as used in the context. According to view 1, speakers can routinely understand a sentence *S* that contains a name the reference of which is fixed by a description, and they can know that *S* expresses a truth, and be prepared to accept and sincerely assent to *S*, but not be in any position to believe, or

even to entertain, the proposition that S semantically expresses. This is puzzling: if one understands S, and S expresses p, then, unless there is something quite special and out of the ordinary about the context, entertaining S should count as entertaining p; and if one comes to accept S, or judge it to be true, then, typically, this should count as believing p, or taking it to be true. If this doesn't, standardly, count as believing the proposition expressed by S, then S will not be of much use in expressing one's beliefs. But if that is so, then it is hard to avoid concluding that S is seriously defective, and perhaps not fully meaningful after all, since it can't be used to perform one of the primary functions of language.

This argument can be strengthened by extending it to the notion of assertion. Someone who believes that the shortest spy is a spy may not believe of anyone that he (or she) is a spy. Similarly, someone who asserts that the shortest spy is a spy may not have said of anyone that he (or she) is a spy. Suppose now that such a speaker goes through a little linguistic ceremony; he uses the description *the shortest spy* to fix the referent of the name *Shifty*—without having any idea who the shortest spy is. If the speaker is successful in introducing the name in this way, then, the sentence *Shifty is a spy* semantically expresses a proposition that can be asserted only by one who says or asserts, of the person who is the shortest spy, that he or she is a spy. But it hardly seems that our speaker would be correctly described as doing this, if he were to assertively utter *Shifty is a spy*. Surely, the speaker wouldn't thereby assert that he or she (demonstrating the individual who in reality is the shortest spy) is a spy. We may suppose that prior to introducing the name, the speaker was not in a position to say anything of this person. But then, going through a formal linguistic ceremony is not going to help. Thus, I believe we must conclude that in assertively uttering *Shifty is a spy*, the speaker does not assert the proposition that, according to view 1, it semantically expresses. The same point could be made about the speaker in Kripke's example who tries to use the description (5) to introduce and semantically fix the referent of PU, which he wrongly takes to designate Princeton University, rather than Kripke's left thumbnail. Just as his utterances of (6b) don't count as expressions of a bizarre belief that the thumbnail is an institution of higher learning, so they don't count as assertions that it is, either.

In my opinion, these results constitute a strong reason for rejecting view 1.<sup>14</sup> If it were correct, then there would be many cases in which

<sup>14</sup> I believe that Kripke now takes the same position, though it is not expressed in *Naming and Necessity*, or in any other of his published writings that I know of.

the referent of a name was semantically fixed by a description for speakers and hearers, even though they did not believe, of the object denoted by the description, that the description applied to it. Where *S* is a sentence containing such a name, we would then have a situation in which *S* means (expresses the proposition) *p*, speakers and hearers understand *S*, but they cannot use *S* to assert *p*, entertain *p*, or express a belief in *p*. That is hard to swallow. If speakers can't use the sentence to assert, entertain, or express a belief in *p*, then it is highly doubtful that the sentence really does mean (semantically express) *p*. All of this argues against the view that we may succeed in introducing a name with a description that semantically fixes its referent, but does not give its content, using any description we like—even one which denotes an object of which we have no independent belief that it satisfies the description. As I see it, this extremely relaxed view about what is required in order to introduce a name by a reference-fixing description should be given up. In its place, we should adopt a view that insists that the introduction of a genuine name by a reference-fixing description (as opposed to the stipulation that the “name” is to be short for the description, or some variant of it) requires us to be in sufficient contact with the object denoted by the description to have independent beliefs about it—including the belief that the description applies to it. With this view in place, we can no longer regard scenario 2 as providing an example of the contingent a priori.

### *Consequences for Kripke's Examples of the A priori*

Let us return, then, to scenario 1. In this scenario, we have seen stick *s* at *t*, and formed an idea of its length, prior to introducing the term *one meter* with the reference-fixing description *the length of stick s at t*. As a result, we have a perceptually justified true belief, of a certain length *l*, that it is the length of stick *s* at *t*. Call this proposition about *l* that we believe on the basis of perceptual evidence *p*. When we introduce the term *one meter* with the reference-fixing description, our knowledge of *p* is part of the knowledge required to understand the sentence *The length of stick s at t is one meter (if s exists at t)*. Since we have this knowledge, we understand the sentence. Moreover, our knowledge of the sentence guarantees not only that we know that it expresses a truth, but also that we grasp the proposition it expresses, and know it to be true. We know, on the basis of the knowledge required to understand the sentence *The length of stick s at t is one meter*,

*if s exists at t*, that the length of stick *s* at *t* is one meter, if *s* exists at *t*. Thus, nothing more than understanding the sentence is required in order for us to know the proposition it expresses. Nevertheless, this knowledge is **not** apriori knowledge, since it rests on our knowledge of *p*, which in turn is justified perceptually. Note, it is not just that we happened to acquire our belief in *p* as a result of perception; *p* is a proposition that can only be known on the basis of such evidence. The same is true of the closely related proposition that the length of stick *s* at *t* is one meter, if *s* exists at *t*. Although it is conceivable that one could know this proposition without knowing the slightly stronger proposition *p*, there is no way one could know it that doesn't rely on empirical evidence for justification. Thus, Kripke's example of the meter stick is **not** a genuine case of the contingent apriori; nor are any of his other examples.<sup>15</sup> Because of requirements on what it takes to successfully introduce a name with a description that semantically fixes its referent, **there are no examples of the contingent apriori which depend upon genuine names the referents of which are semantically fixed by descriptions.**

### Are There Examples of the Contingent Apriori?

I believe there are, the best examples being those that involve the actuality operator discussed in chapter 14.<sup>16</sup> Briefly put, this operator works as follows: Let *S* be a sentence that expresses a proposition *p* in a context of utterance *C*. Then, in *C*, **Actually** *S* expresses a proposition that says, of the actual world-state of *C*, that it is a world-state in which *p* is true—i.e., it says that *p* is true with respect to (or given) the way the world actually is. Hence, the proposition expressed by **Actually** *S* in *C* is true when evaluated with respect to an arbitrary possible world-state *w* iff *S* is true with respect to the actual world-state of *C*; since this never changes—i.e., since whether or not *S* is true with respect to the fixed actual world-state never changes as one moves from

<sup>15</sup> Nathan Salmon reports that Kripke told him, in personal communication, that in his discussion of the meter-stick example in *Naming and Necessity* he had in mind “a case in which the reference-fixer sees *S* there in front of him and uses the description referentially to refer to that length” (p. 200 of Salmon’s “How to Measure the Standard Metre”). Nevertheless, Kripke did not conclude, as Salmon and I do, that in that case the example cannot properly be regarded as an instance of the contingent apriori.

<sup>16</sup> Examples of this type were pointed out by David Kaplan in “On The Logic of Demonstratives,” *Journal of Philosophical Logic* 8 (1979): 81–98, and in “Demonstratives.”

one merely possible world-state to another—whenever *S* expresses a truth in *C*, *Actually S* expresses a necessary truth. The corresponding fact about descriptions is that whenever *the x: Fx* successfully denotes a unique individual *o* in the actual world-state of a context of utterance, *the x: actually Fx*, as used in that context, denotes *o* with respect to all possible world-states in which *o* exists, and never denotes anything else. Hence *actually* is a rigidifier which, when added to a description that denotes a unique individual with respect to the context of utterance, results in a new description that rigidly denotes that individual with respect to all possible world-states in which that individual exists. The description *the x: actually Fx* may also be paraphrased *the actual F*. If asked what this description means, we answer that it may be paraphrased *the unique object which is F in @*, where '@' designates the way that the world actually is—i.e., the maximally complete property that the world actually instantiates. Thus, a sentence *The actual F is G* says, in effect, that which is said by *The unique object which is F in @ is also G*.

With this in mind, we may consider the following examples.

- 8a. Princeton University has a philosophy department iff actually, Princeton University has a philosophy department.
- b. If stick *s* exists at *t* (and hence has a unique length at *t*), then the length of stick *s* at *t* is the actual length of stick *s* at *t*.

These are genuine examples of the contingent apriori. First contingency. Since Princeton has a philosophy department, the right-hand side of the biconditional (8a) is a necessary truth and the left-hand side is a contingent truth; hence (8a) is itself contingent. (8b) is also contingent, since although it is true with respect to the actual state of the world, it is false with respect to possible states of the world *w* in which *s* exists at *t*, but either its length at *t* is greater than the length it has with respect to the actual state of the world, due to its having been heated and stretched in *w*, or its length at *t* is less than the length it has with respect to the actual state of the world, due to its having been cooled and compressed in *w*.

Next aprioricity. The basic principle to which we need to appeal is (9a), and its corollary (9b).

- 9a. For any proposition *p* and possible world-state *w*, one may know *p* in *w* on the basis of evidence *e* iff in *w*, one may know, of *w*, that it is a world-state in which *p* is true, on the basis of that same evidence *e*.

- b. For any proposition  $p$  and possible world-state  $w$ , one may know  $p$  apriori in  $w$  iff in  $w$ , one may know apriori, of  $w$ , that it is a world-state with respect to which  $p$  is true.

Taking these principles plus certain obvious subordinate principles as given, we can establish that the proposition expressed by (8a) is knowable apriori as follows.<sup>17</sup>

- 10(i) The proposition that Princeton University has a philosophy department iff Princeton University has a philosophy department is knowable apriori (in the actual world-state). Call this *the proposition that  $P$  iff  $P$* .
- (ii) From (i) plus (9b), it follows that in the actual world-state @, it is knowable apriori, of @, that it is a world-state with respect to which the proposition that  $P$  iff  $P$  is true—i.e., that it is a world-state  $w$  which is such that the proposition that  $P$  iff  $P$  is true with respect to  $w$ .
- (iii) So, in (or with respect to) @, it is knowable apriori, of @, that it is a world-state  $w$  which is such that the proposition that  $P$  is true with respect to  $w$  iff the proposition that  $P$  is true with respect to  $w$ .
- (iv) So, in (with respect to) @, it is knowable apriori, of @, that it is a world-state  $w$  which is such that the proposition that  $P$  is true with respect to  $w$  iff the proposition that  $P$  is true with respect to @.
- (v) So, in (with respect to) @, it is knowable apriori, of @, that it is a world-state  $w$  which is such that the proposition that  $P$  iff the proposition that  $P$  is true with respect to @ is true with respect to  $w$ —i.e., in @, it is knowable apriori, of @, that it is a world-state with respect to which the proposition that  $P$  iff the proposition that  $P$  is true in @ is true.
- (vi) From (v) and (9b) it follows that, in (with respect to) @, the proposition that  $P$  iff the proposition that  $P$  is true in @ is knowable apriori.

<sup>17</sup> In (10), I use ‘ $P$ ’ as an abbreviation of the sentence ‘Princeton University has a philosophy department’.

- (vii) Since @ is the actual world, the proposition that the proposition that P is true in @ is the proposition expressed by *Actually P*. Thus, it follows from (vi) that the proposition expressed by *P iff actually P*—i.e., the proposition expressed by (8a)—is knowable apriori (in the actual world).

It may be helpful to think of this argument schematically, along the lines of (11).<sup>18</sup>

- 11(i) P iff P  
 (ii) So,  $\lambda w[(P \text{ iff } P) w]$  @  
 (iii) So,  $\lambda w [P(w) \text{ iff } P(w)]$  @  
 (iv) So,  $\lambda w [P(w) \text{ iff } P(@)]$  @  
 (v) So,  $\lambda w [(P \text{ iff } P@)w]$  @  
 (vi) So, P iff P@  
 (vii) So, P iff actually P

In (i) of (11), we start with a claim that is obviously knowable apriori. We then use (9b) to assure us that (ii) is apriori, where (ii) tells us that @ has the property of being a world-state  $w$  with respect to which it is true that P iff P. (iii) is an apriori consequence of (ii); it tells us that @ has the property of being a world-state with respect to which the proposition that P is true iff it is a world-state with respect to which the proposition that P is true. The key move is to (iv).<sup>19</sup> Consider an analogy. Suppose that  $m$  has the property that any man has just in case he loves Anna iff he loves Betty—i.e.,  $\lambda x[Lxa \text{ iff } Lxb]$   $m$ . Then  $m$  also has the property that any man has just in case he loves Anna iff  $m$  loves

<sup>18</sup> In (11) I continue to use ‘P’ as an abbreviation for ‘Princeton University has a philosophy department’. I use ‘@’ to designate the actual state of the world, and I use ‘ $w$ ’ as a variable over world-states. When  $S$  is any sentence, I use  $S@$  to express the claim that attributes the property of being true with respect to the actual world-state to the proposition that  $S$  expresses, and I use the formula  $S(w)$  to express the claim that that proposition is true with respect to the world-state which is the value of ‘ $w$ ’. Where  $\dots v \dots$  is any formula containing one or more free occurrences of the variable  $v$ ,  $\lambda v[\dots v \dots]$  is a predicate expressing a property that is true of an object  $o$  iff  $\dots v \dots$  is true of  $o$  relative to an assignment of  $o$  to ‘ $v$ ’.

<sup>19</sup> The idea for this move can be found on pp. 210–11 of Gareth Evans, “Reference and Contingency,” in his *Collected Papers* (Oxford: Oxford University Press, 1985). Despite this insight, in my opinion Evans misconstrues examples involving names the referents of which are fixed by description. See pp. 148–50 of my review of his *Collected Papers*, *Journal of Philosophy* 86 (1989).

Betty—i.e.,  $\lambda x[Lx \text{ iff } Lmb] m$ . Similarly, (iii) tells us that @ has the property that any world-state has just in case the proposition that P is true with respect to that world-state iff the proposition that P is true with respect to that world-state—i.e.,  $\lambda w [P(w) \text{ iff } P(w)] @$ . It follows that @ also has the property that any world-state has just in case the proposition that P is true with respect to that world-state iff the proposition that P is true with respect to @—i.e.,  $\lambda w [P(w) \text{ iff } P(@)] @$ . (v) is an apriori consequence of (iv); it tells us that @ has the property of being a world-state with respect to which the proposition that P iff P is true with respect to @ is true. We then use (9b) again (this time in the right-to-left direction) to assure us that the proposition expressed by (vi) is knowable apriori, which is the same as the proposition expressed by (vii)—namely, the proposition expressed by (8a).<sup>20</sup>

Thus, in addition to being contingent, the proposition expressed by (8a) is knowable apriori. A similar demonstration can be given that the proposition expressed by (8b) is knowable apriori, on the basis of the obvious fact that the proposition expressed by (8c) is knowable apriori.

8c. If stick *s* exists at *t* (and hence has a unique length at *t*), then the length of stick *s* at *t* is the length of stick *s* at *t*.

Since the proposition expressed by (8b) is also contingent, it too is a genuine example of the contingent apriori. In general, for every

<sup>20</sup> There is a potential objection to principles (9a) and (9b) (in the right-to-left direction) which suggests that they may not be true strictly as stated. However, the objection does not affect our use of them in the above derivation. The objection is based on the following scenario. Suppose one describes a possible world-state *w* to oneself in some detail, and concludes, correctly, that it is a possible world-state with respect to which a certain proposition *q* is true. This may count as knowledge, of *w*, that *p* is true with respect to it. Let us suppose that it does. Suppose further that *w* is, in fact, the actual world-state—the world-state with respect to which one is doing the relevant describing and thinking—but that one doesn't realize that the world-state one has been describing and thinking about is the state that the world is actually in. In such a case, one might believe of the actual world-state that *q* is true with respect to it, without believing *q*, and without having any justification for believing *q*. If there are possible scenarios like this, then they threaten the right-to-left direction of (9a), and perhaps even (9b). However, even if they falsify (9a) and (9b) as stated, they do not invalidate the above derivation, and in particular, they do not invalidate the move from (v) to (vi). In the scenario relevant to the derivation, the agent starts with apriori knowledge of a certain proposition, and moves to knowledge of the actual world-state that the proposition is true with respect to it—which he can express using the demonstrative *this world-state* with the understanding that it refers to the world-state he directly experiences. Throughout the entire derivation, where we use '@' the agent can use *this world-state* with that understanding. If the agent does this, then all the argumentative moves, including the move from (v) to (vi), will be transparent.



Kripke-style example of a sentence . . . n . . . containing a name *n* the referent of which is fixed by a description *the D* that is supposed to be an instance of the contingent apriori (but really isn't), the corresponding sentence in which *the actual D* is substituted for *n* is a genuine example of the contingent apriori—i.e., it is a sentence that expresses a proposition which, though true, and knowable apriori, is not a necessary truth.<sup>21</sup> In short, although none of Kripke's putative examples of the contingent apriori are correct, and although names the referents of which are semantically fixed by descriptions in his sense **never** give rise to genuine instances of the contingent apriori, Kripke showed us what to look for, and put us on the track to finding the examples that really do make his point. Surely, that is the thing to remember.

<sup>21</sup> Here, I rely on our earlier discussion in chapter 14 which supports the claim that for Kripke, even when a name has its referent semantically fixed by a description *the D*, it is not synonymous with *the actual D*.

# CHAPTER 17

## NATURAL KIND TERMS AND THEORETICAL IDENTIFICATION STATEMENTS

### CHAPTER OUTLINE

1. *Extending the Analysis of Names to Natural Kind Terms*  
Proper names and natural kind predicates: similarities and differences  
Rigid designation is not defined for predicates
2. *Identity Sentences Containing Natural Kind Predicates*  
The logical forms of such sentences
3. *The Modal Status of Theoretical Identities Involving Simple Natural Kind Predicates*  
Non-descriptiveness and the explanation of why some of these sentences are necessary, if true
4. *Extending the Account*  
The necessity of certain theoretical identities involving semantically complex, theoretically interesting predicates
5. *The Epistemic Status of Theoretical Identities Involving Natural Kind Predicates*  
Why some are knowable only a posteriori
6. *A Final Challenge to the Necessary A Posteriori*  
The hypothetical objector and Kripke's unsatisfying reply
7. *A Proper Response to the Challenge: Essential Properties, Impossible World-States, and the Necessary A Posteriori*  
The difference between epistemic and metaphysical possibility

### **Extending the Analysis of Names to Natural Kind Terms**

In this chapter, we will discuss how Kripke extends his central theses about proper names to the broader, more heterogeneous, and more philosophically significant class of natural kind terms. His claim that true theoretical identity sentences involving such terms are examples of the necessary a posteriori is particularly important in this connection.

This claim captured the imagination of philosophers, and held out the promise of important and widespread applications in philosophy as a whole, in a way that went well beyond the perceived implications of his similar claim about identity sentences containing proper names. Still, his analysis of proper names laid the foundation for his treatment of natural kind terms, and it is important to understand how his views about the latter grew out of his views about the former. The central theses about names defended in the first two lectures of *Naming and Necessity* are summarized by T1–T4.

THESES ABOUT PROPER NAMES

- T1. Proper names are non-descriptonal: (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 with respect to *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.
- T2. The referent of a proper name is initially fixed 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 fixed usually doesn't matter. Typically, speakers further 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.
- T3. Proper names are rigid designators—i.e., 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.
- T4. Identity sentences involving different but coreferential names (or other rigid designators) express necessary truths. Nevertheless, often these truths are knowable only *aposteriori*.

In lecture 3 of *Naming and Necessity*, 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 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 established normally doesn't matter—just as with proper names. As a result, speakers further down the 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 (or identification) sentences that are necessary but aposteriori. Examples of such theoretical identity sentences specifically discussed in *Naming and Necessity* are:

*Water is H<sub>2</sub>O. Flashes of lightning are flashes of electricity. Light is a stream of photons. Gold is the element with atomic number 79. Cats are animals. Whales are mammals. Heat is the motion of molecules.*

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 rigid designation, 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."<sup>1</sup> However, there is a potential difficulty here that has not

<sup>1</sup> *Naming and Necessity*, p. 140, my emphasis.

been widely appreciated. Kripke gives no separate definition of what it means to say that a natural kind term is rigid; nor does he provide distinct arguments to show that they 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.

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.<sup>2</sup>

A little later, summing up his views, he 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’.<sup>3</sup>

It appears from these passages that Kripke intends his general theses about natural kind terms to apply, at least in some form, to terms of various syntactic and semantic categories. This raises obvious questions, among them, “What is it for a predicate to be a rigid designator?”, “Are natural kind predicates involving general terms like *cow*, *tiger*, *animal*, *chunk of gold*, and *flash of lightning* rigid?”, “What is it for a sentence containing a pair of natural kind predicates to count as a theoretical identity sentence?”, and “Are such sentences necessary if true? If so, do some of them express claims that are knowable only

<sup>2</sup> Ibid., p. 127, my emphasis.

<sup>3</sup> Ibid., p. 134.

aposteriori?" Since Kripke nowhere explicitly answers these questions, it is a matter of speculation how he would deal with them.

Although these questions are certainly important, it might seem as if they can easily be answered. In particular, it might seem straightforward how to define rigidity for predicates. Think again about rigidity for singular terms. In the case of a singular term, *t*, the object designated by *t* with respect to a world-state *w* is the extension of *t* with respect to *w*, and the claim that *t* is rigid entails that it has the same extension with respect to every world-state at which it has an extension at all. In the case of a predicate, *P*, the extension of *P* with respect to a world-state *w* is the set of objects that *P* applies to, or is true of, with respect to *w*. One idea for characterizing a notion of rigidity for predicates parallel to the notion of rigidity for singular terms would be to stipulate that *P* is rigid only if *P* has the same extension with respect to any world-state at which it has a non-null extension. However, this clearly won't do. Consider, for example, the natural kind predicate *is an animal*. Its extension with respect to a world-state is the set of all things that are animals at that world-state. Since the set of individual animals varies from one world-state to the next, this definition of rigidity would classify it as non-rigid. Since the same point could be made for virtually every predicate of contingently existing objects, rigidity for predicates shouldn't be defined as requiring sameness of extension at different world-states (if it is to be a theoretically significant notion).

Nor will it do to say that a predicate is rigid iff there is a unique property which it expresses that determines its extension at each possible world. It could, of course, be argued that there is such a property in the case of natural kind predicates involving terms like *cow* and *animal*—namely, the property of being a cow and the property of being an animal. However, the same could be said for virtually any predicate. For example, the predicates *is a bachelor* and *is a philosopher* also express properties that determine their extensions with respect to arbitrary possible world-states. In general, for virtually any predicate *F* one can think of, and any world-state *w*, the extension of *F* with respect to *w* is the set of things that have, at *w*, the property expressed by *being an F*. But there is no point in defining a notion of rigidity for predicates according to which virtually all predicates turn out, trivially, to be rigid.

There is, it should be noted, an obvious alternative definition which does not have this consequence, and which is a natural extension of Kripke's definition of rigidity for singular terms. The idea is that a

predicate is rigid iff it is an essentialist predicate, where the latter is characterized as follows:

EP. A predicate *P* is essentialist iff for all possible worlds *w* and objects *o*, if *P* applies to *o* with respect to *w*, then *P* applies to *o* in all worlds in which *o* exists.

Two other ways of expressing the same idea are:

EPa. A predicate *P* (of English) is essentialist iff *Necessarily any individual that is (was) P could not have existed without being P* expresses a truth.

EPb. A predicate *P* is essentialist iff the property it expresses is an essential property of anything that has it.

The parallels between this interpretation of rigidity for predicates and the corresponding theses for rigid singular terms are obvious. For example, the linguistic test for the essentiality/rigidity of English predicates provided by EPa is similar to Kripke's rigidity test for singular terms.

RT. A singular term *t* (of English) is a rigid designator of an object iff *The individual that is (was) t could not have existed without being t (and no one other than that individual could have been t)* expresses a truth.

According to this test, the name *Aristotle* is rigid, since necessarily, the individual who was Aristotle could not have existed without being Aristotle (and no one else could have been Aristotle); whereas the definite description *the teacher of Alexander* is not rigid. Similarly, the predicate *is an animal* is (arguably) essentialist, and therefore, on this interpretation, rigid, since necessarily anything that is an animal could not have existed without being an animal. By contrast, *is a philosopher* is not essentialist/rigid.

Although this definition of rigidity for predicates is more plausible than the other two, it also won't work. As I demonstrate in chapter 9 of *Beyond Rigidity*, it suffers from two crushing defects. First, not all natural kind predicates covered by Kripke's theory are essentialist. Since these predicates occur in necessary a posteriori statements of theoretical identity, the presence of essentialist predicates is **not** crucial to Kripke's explanation of these cases.<sup>4</sup> Second, even though many natural kind

<sup>4</sup> The example (10b) discussed below, which contains the two-place natural kind predicate *hotter than*, is an illustration of this point.

predicates that Kripke discusses are essentialist, the fact that they are essentialist does not explain the necessity of true statements of theoretical identity involving them.<sup>5</sup> Since it is precisely an explanation of such cases of the necessary aposteriori that we are most interested in, extending the notion of rigid designation to predicates in this way is of no help.

At this point, there are two different ways in which we might proceed. One way would be to ignore predicates and restrict ourselves to those uses of natural kind terms in which they can naturally be understood to be functioning as names of natural kinds themselves—where each kind is taken to be an abstract object which has concrete individuals as instances (or members).<sup>6</sup> On this approach, we would appeal to the concept of rigidity as defined for singular terms, and try to apply Kripke's doctrines about identity sentences of the form  $\alpha = \beta$  to reach results about necessary aposteriori identity statements formulated using names of abstract kinds.<sup>7</sup> A different way of proceeding would be to focus on natural kind predicates directly, to put aside the concept of rigidity, and to identify other ways in which, according to Kripke, natural kind predicates are similar to names. On this approach, we first have to decide what counts as a theoretical identity sentence involving natural kind predicates; next, we have to identify features of these predicates which they share with proper names that are responsible for (i) the fact that true theoretical identification sentences involving them are necessary, if true, and (ii) the fact that sometimes these sentences express claims that are knowable only aposteriori.

There are two reasons for favoring the second approach, which we will adopt here. First, it is more general; for some natural kind terms,

<sup>5</sup> See pp. 257–59 of *Beyond Rigidity*.

<sup>6</sup> These uses are explained in chapter 11 of *Beyond Rigidity*.

<sup>7</sup> There is a different way of thinking of some natural kinds which takes them to be large, scattered concrete objects. For example, gold is sometimes thought to be a scattered object—the gold-fusion—of which each individual gold thing is a part. However, if 'gold' is treated as a singular term naming this unusual object, then it is hard to think of it as a rigid designator—for it would seem that gold could have existed, even if none of the chunks of matter that are gold in the world as it actually is existed. Let  $w$  be a possible world-state with respect to which that is so. With respect to  $w$ , the proposal that 'gold' is a singular term that designates the sum of all existing gold in the world would have to treat 'gold' as designating a gold-fusion with no parts in common with the gold-fusion it designates with respect to the actual state of the world. Since it is hard to see how the gold-fusion in  $w$  could be the very same concrete object as the gold-fusion in the world as it actually is, it is hard to see how, on this proposal, 'gold' could be a rigid designator. I will therefore put this possible interpretation aside. For further discussion, see chapter 11, footnote 19 of *Beyond Rigidity*; see also Kathrin Koslicki, "The Semantics of Mass Predicates," *Noûs* 38 (1999):46–91.



e.g., *star* and *electron*, their only natural use is as predicates rather than names, and it is awkward and artificial to try to come up with complex singular terms appropriately related to them that can be used to play the naming role required by the first approach. Second, in the case of many natural kind terms—e.g., the adjective *red*—which are capable of functioning both as predicates of concrete individuals (*The ball is red*) and as names of abstract kinds (*Red is a color*), we seem to understand the predicate prior to understanding the associated singular term, which suggests that the semantic properties of predicates should be specifiable without having to appeal to the related singular term.

### Identity Sentences Containing Natural Kind Predicates<sup>8</sup>

In order to understand and evaluate the doctrine that theoretical identity sentences containing natural kind predicates are necessary, if true, we must first understand what sentences we are talking about. Normally, when we talk about identity sentences, we have in mind sentences of the form (1), where  $\alpha$  and  $\beta$  are names or other singular terms.

$$1. \alpha = \beta$$

If (1) is true, then these singular terms refer to the same thing. This means that if  $\alpha$  and  $\beta$  are rigid, they will refer to the same thing with respect to every possible world-state in which the (actual) referent of  $\alpha$  (and  $\beta$ ) exists. So, if (1) is true and the terms are rigid, then (1) is true with respect to all possible world-states, and hence is necessary.<sup>9</sup> However, when natural kind predicates are involved, the question of what sort of sentence is to count as an identity is more complicated. At various points throughout lecture 3, Kripke discusses a class of statements involving natural kind terms that he calls “theoretical identifications.” The initial examples of these sentences are (2–5).<sup>10</sup>

2. Light is a stream of photons.
3. Water is H<sub>2</sub>O.
4. Lightning is an electrical discharge.
5. Gold is the element with the atomic number 79.

<sup>8</sup> The material in this section is also discussed in chapter 9 of *Beyond Rigidity*.

<sup>9</sup> For simplicity, I ignore possible world-states in which the referent of the terms doesn't exist.

<sup>10</sup> *Naming and Necessity*, p. 116.

It is worth noting that (2) and (4) do not even appear to be sentences in which a pair of singular terms flank the identity sign.

Later in the lecture, Kripke considers other statements that he seems to place in the same category as the original examples (2–5). One such statement is (6).

6. Cats are animals.

Clearly, this is not a sentence the logical form of which is given by (1). Rather, its logical form is normally regarded by philosophers as involving a pair of predicates, and as being representable (at least for many purposes) by something like (7).

7.  $\forall x (x \text{ is a cat} \supset x \text{ is an animal})$

For all  $x$ , if  $x$  is a cat, then  $x$  is an animal.

Although it doesn't contain the identity predicate, (7) might broadly be classified as an identity statement on the grounds that it identifies each cat with some animal. In general, English sentences of the form

8a. (All)  $A$ 's are  $B$ 's, An  $A$  is a  $B$  (on one of its uses)

and

8b. All and only  $A$ 's are  $B$ 's, Something is an  $A$  iff it is  $B$  (on one use)

may naturally be counted as statements that identify  $A$ 's with  $B$ 's, even though they are routinely represented by formulas of the form (9), which do not contain the identity sign.

9a.  $\forall x (Ax \supset Bx)$

b.  $\forall x (Ax \leftrightarrow Bx)$

Next consider (4), from Kripke's original list of theoretical identifications. Here, the expression *is an electrical discharge* seems to function as a predicate, just as the expression *is a cat* does in the sentence *Felix is a cat*. Thus (4) is not the sort of identity statement in which a pair of singular terms flanks the identity sign. Rather, it says that any individual instance of lightning is an electrical discharge, in which case it is to be understood on the model of (9a). Once this is recognized, one might analyze (2) as also being of this form.

The lesson here is that theoretical identity statements need not involve singular terms, or the identity sign, but instead may have the

form of universally quantified conditionals or biconditionals. This point is further illustrated by the following passage from *Naming and Necessity*.

[I]t turns out that a **material object is (pure) gold if and only if the only element contained therein is that with atomic number 79**. Here, the ‘if and only if’ can be taken to be strict (necessary). In general, science attempts, by investigating basic structural traits, to find the nature, and thus the essence (in the philosophical sense) of the kind. The case of natural phenomena is similar; such theoretical identifications as ‘heat is molecular motion’ are necessary, though not apriori. The type of property identity used in science seems to be associated with necessity, not with apriority, or analyticity: **For all bodies x and y, x is hotter than y if and only if x has a higher mean molecular kinetic energy than y**. Here the coextensiveness of the predicates is necessary, but not a priori.<sup>11</sup>

In this passage, Kripke gives two examples of necessary aposteriori statements of theoretical identification that have the form of universally quantified biconditionals. Moreover, he seems to suggest that the doubly quantified biconditional (10b) may be an analysis of the identity (10a).

10a. Heat is molecular motion.

b. For all bodies x and y, x is hotter than y if and only if x has a higher mean molecular kinetic energy than y.

In light of all this, it seems clear that Kripke’s claims about theoretical identity sentences involving natural kind terms should be seen as encompassing cases in which the terms function as predicates, and the identity sentences have analyses which are universally quantified conditionals or biconditionals.

### **The Modal Status of Theoretical Identities Involving Simple Natural Kind Predicates<sup>12</sup>**

What is it about natural kind predicates that makes Kripke confident that theoretical identity sentences containing them are (often) necessary,

<sup>11</sup> Ibid., p. 138, my emphasis.

<sup>12</sup> The material in this section and the next draws heavily on the first half of chapter 10 of *Beyond Rigidity*.

if true? On page 140 he seems to tell us when he announces that “theoretical identities, according to the conception I advocate, are generally identities involving two rigid designators and therefore are examples of the necessary *a posteriori*.” However, I have already noted that this answer is problematic and insufficiently general, since the notion of a rigid designator has not been defined for predicates. Fortunately, Kripke has a better answer available. Like names, natural kind predicates are non-descriptonal—i.e., they are not synonymous with descriptions associated with them by speakers, and their reference (the set of things they correctly apply to) is not semantically determined by such descriptions. This, I suggest, is the important parallel with proper names.

In the first section of this chapter, I quoted passages from pages 127 and 134 which indicate that Kripke intended his analysis of the parallels between proper names and natural kind terms to apply to kind terms of several grammatical categories, including those that function as predicates. It is striking that, although he sometimes mentioned rigidity as being included among the parallels between names and natural kind terms, it is non-descriptonality that occupies center stage in the discussion immediately following the second of these two passages.

[M]y 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’.

Mill, as I have recalled, held that although some ‘singular names’, the definite descriptions, have both denotation and connotation, others, the genuine proper names, had denotation but not connotation. Mill further maintained that ‘general names’, or general terms, had connotation. Such terms as ‘cow’ or ‘human’ are defined by the conjunction of certain properties which pick out their extension—a human being, for example, is a rational animal with certain physical characteristics. The hoary tradition of definition by *genus* and *differentia* is of a piece with such a conception. If Kant did, indeed, suppose that ‘gold’ could be *defined* as ‘yellow metal’, it may well be this tradition which led him to the definition. . . .

The modern logical tradition, as represented by Frege and Russell, disputed Mill on the issue of singular names, but endorsed him on that of general names. Thus *all* terms, both singular and general, have a ‘connotation’ or Fregean sense. More recent theorists have followed Frege and Russell, modifying their views only by replacing the notion of a sense as given by a particular conjunction of properties with that of a sense as given by a ‘cluster’ of properties, only *enough* of which need apply. The present view, directly reversing Frege and Russell, (more or less) *endorses* Mill’s view of *singular* terms, but *disputes* his view of *general* terms.<sup>13</sup>

This passage occurs in the middle of the third lecture, where Kripke says he is “recapitulating” his main theses about natural kind terms before going on to a discussion of the mind-body identity theory. His recapitulation involves the point just made about the non-descriptiveness of natural kind terms (of various grammatical categories), plus his account of how the reference of these terms is fixed.

It is evident that for Kripke the objects to which natural kind terms correctly apply with respect to different possible states of the world are not determined by whatever descriptive properties, if any, speakers associate with the terms. It is true, of course, that speakers often associate them with descriptive properties that they use to identify particular instances of the kinds. However, these properties typically fail to provide necessary and sufficient conditions for something to be a member of a kind (with respect to a world-state), and sometimes the properties associated with such a term are not even true of actual instances of the kind (as when speakers think of a whale as a kind of fish). Just as someone can successfully use a proper name to refer to an object without associating the name with descriptive properties that uniquely apply to its referent, so a speaker can successfully use a natural kind predicate to say something about the members of a kind, even if he lacks the ability to accurately describe either the kind itself or its instances.

Next, we turn to Kripke’s positive account of how the extension of a natural kind predicate (the class of things to which it correctly applies) with respect to an arbitrary world-state is semantically determined. According to that account, the predicate is first associated by speakers with a kind—either ostensively or via a description. In the ostensive case, speakers directly associate the predicate with a sample of individuals, which they presume to be instances of a single natural kind *k* of a given

<sup>13</sup> *Naming and Necessity*, pp. 134–35.

type (e.g., a single substance, or a single species). Since kinds are hierarchically arranged in terms of genus and species, there is never a **single** natural kind of which members of the sample are instances. Thus, the presumption that  $k$  be of some definite type is crucial when a natural kind term is introduced in this way. (For example, when we introduce *gold* ostensively in terms of a sample, we think of the kind as being a type of metal.) In the case in which a predicate is associated with a kind descriptively, speakers employ a description that picks out a unique kind, often by appeal to contingent properties of the kind or its instances. Once the kind  $k$  has been determined, either ostensively or by description, it is understood that for any world-state  $w$ , the extension of the predicate with respect to  $w$  is to be the set of instances of  $k$  with respect to  $w$ .<sup>14</sup>

With this in mind, let us return to the question of whether theoretical identity sentences involving natural kind predicates are necessary, if true. The explanation that follows, though constructed from central aspects of Kripke's account, is not explicit in his text. Moreover, it does not cover every "theoretical identification sentence" mentioned there. However, it does apply to many of the most important examples. Since I don't know of any other explanation that covers all of his examples, and since the examples themselves seem to fall rather naturally into different categories, I do not regard the incompleteness of the explanation that follows to be a crippling defect.<sup>15</sup>

The explanation concerns theoretical identity sentences of the form (8a,b), which are analyzed along the lines of (9a,b).

8a. (All) A's are B's, An A is a B (on one of its uses)

8b. All and only A's are B's, Something is an A iff it is B (on one use)

9a.  $\forall x (Ax \supset Bx)$

b.  $\forall x (Ax \leftrightarrow Bx)$

<sup>14</sup> This picture should not be credited to Kripke alone. Substantially the same account was developed independently by Hilary Putnam. See in particular Putnam's "Is Semantics Possible?" in H. Kiefer and M. Munitz, eds., *Language, Belief and Metaphysics* (Albany: State University of New York Press, 1970); "Explanation and Reference," in G. Pearce and P. Maynard, eds., *Conceptual Change* (Dordrecht: Reidel, 1973); "Meaning and Reference," *Journal of Philosophy* 70 (1973):699-711; "The Meaning of 'Meaning'," in K. Gunderson, ed., *Language, Mind, and Knowledge*, Minnesota Studies in the Philosophy of Science, no. 7 (Minneapolis: University of Minnesota Press, 1975). All of these papers are reprinted in Hilary Putnam, *Philosophical Papers*, vols. 1 and 2 (Cambridge: Cambridge University Press, 1975).

<sup>15</sup> For further discussion of the scope of this explanation, as well as the relationship between it and Kripke's text, see chapters 9 and 10 of *Beyond Rigidity*.

To begin with, I will restrict attention to those sentences of this form in which the predicates A and B are different, semantically simple, natural kind predicates (often, single words). In addition, it will be convenient, initially, to further restrict the case by assuming that the reference of A is fixed ostensively. That is, we assume that the reference of A is fixed by stipulating that it is to apply to all and only instances of the unique natural kind of a certain type T (e.g., a substance or species) of which nearly all elements of the sample used in introducing it are instances. Unlike A, the predicate B is not required to be an ostensive natural kind term; rather, we allow its extension to be fixed either ostensively or by description. However, we do require that A and B be predicates associated with kinds of the same type T—for instance, both may be substance terms, both may be species terms, or both may be of some other category.<sup>16</sup> When A and B are related in this way, it may turn out that the relationship between their extensions is not evident to competent speakers, but is discoverable only by empirical investigation. Let us suppose that this is so, and that, as the result of empirical investigation, it has now been discovered that every object in the extension of A is also in the extension of B. This discovery establishes that sentence (8a) is true. What needs to be shown is how this apparently modest result, combined with the semantic character of the predicates, suffices to guarantee that they designate the same natural kind, and hence to guarantee that (8a) is necessary (and (8b) is too).

To show this, we reason as follows: (i) From the assumption that the ostensive natural kind predicate A has successfully been introduced, it follows that there is a **unique** natural kind  $k_A$  (of a given type T) of which nearly all members of the sample associated with A are instances, and A applies (with respect to a world-state) to all and only instances of  $k_A$  (with respect to that world-state). (ii) From the assumption that the natural kind predicate B has successfully been introduced, it follows that there is a natural kind  $k_B$  which is such that B applies (with respect to a world-state) to all and only members of  $k_B$  (with respect to that

<sup>16</sup> This requirement imposes an important limitation on the explanation being proposed—namely, that it does not apply to examples like *Cats are animals* or *Whales are mammals*. In these cases, although the A and B predicates are natural kind terms, they are not intended to designate kind of the same type; *cats* and *whales* are species terms, whereas *animals* and *mammals* designate kinds of a higher type. Thus, if, as seems plausible, these sentences are necessary if true, some different and independent explanation must be given. The most likely explanation is that it is a metaphysically essential property of the pair consisting of the species cat and the kind animal that any instance of the former is an instance of the latter; ditto for the pair consisting of the species whale and the kind mammal.

world-state). (iii) By hypothesis, the two predicates designate kinds of the same type (type T); thus both  $k_A$  and  $k_B$  are species, or both are substances, or both are kinds of some other category. (iv) If the theoretical identity sentence (8a) is true, then (since nearly all objects in the A-sample are A's) nearly all of the objects in the A-sample are B's, and hence they are instances of kind  $k_B$  as well as kind  $k_A$ . (v) Since, by hypothesis, the A-sample determines a **single** kind (of the given type T—a single species, substance, etc.), it follows that kind  $k_A =$  kind  $k_B$ . (vi) But this means that in addition to (8a), (8b) must also be true. (vii) Moreover, both must be **necessary truths**, since from steps (i), (ii), and (v) it follows that for all world-states  $w$ , the extension of A with respect to  $w =$  the set of instances of  $k_A$  with respect to  $w =$  the set of instances of  $k_B$  with respect to  $w =$  the extension of B with respect to  $w$ . (viii) So, if A and B are natural kind predicates of the sorts indicated, then a theoretical identity sentence (8b) involving those predicates is a necessary truth, if the corresponding sentence (8a) is true.

This provides us with the paradigm (11).

11a.  $\forall x (Ax \supset Bx)$  is true, when understood as intended.

b. A and B are simple natural kind predicates of the same type (e.g., both species predicates, both substance predicates, etc.). Moreover, A “designates” the unique natural kind (of the specified type) instantiated by nearly all members of its associated sample. Where  $k_A$  and  $k_B$  are the kinds associated with (“designated by”) A and B, respectively, the extensions of A and B with respect to a world-state  $w$  are the sets of individuals that are members, with respect to  $w$ , of  $k_A$  and  $k_B$ , respectively.

c.  $\forall x (Ax \supset Bx)$  and  $\forall x (Ax \leftrightarrow Bx)$  are both necessary truths.

Although the argument may initially seem surprising, it is easy to see why it works. On Kripke's account, the linguistic mechanism that fixes the extension of ostensive natural kind predicates guarantees that, if the term has been introduced successfully, then there can be **only one** natural kind, of the appropriate sort, of which nearly all members of the sample are instances. So, when A is ostensive, and the predicate B is related to A in the appropriate way, finding out that everything in the extension of A is in the extension of B is tantamount to finding out that the kind associated with A is identical with the kind associated with B. This powerful conclusion does not come from nowhere.



Rather, it comes from the fact that the semantic presuppositions associated with the introduction of the natural kind predicates include very substantial non-linguistic claims—including the claim that the individuals in the sample associated with the ostensive predicate A are members of a single natural kind of a given type, and the claim that the predicates A and B stand for kinds of the same type.<sup>17</sup> Because of this, one should be careful about drawing the moral from this example that it is easy to know or establish certain necessary a posteriori truths. One could equally well conclude that it is rather daunting to establish certain non-modal identification statements.<sup>18</sup>

### Extending the Account

So far, we have restricted paradigm (11) to universally quantified conditionals and biconditionals involving semantically simple natural kind predicates. However, we know that Kripke was willing to characterize a wider class of theoretical identity sentences, including (12) and (13), as being necessary, if true.

12. For all bodies x and y, x is hotter than y iff x has a higher mean molecular kinetic energy than y.
13. For all x, x is a drop of water iff x is a drop of a substance molecules of which contain two hydrogen atoms and one oxygen atom.

Example (13) is a theoretical identification sentence of the form (9b) in which one of the predicates, *is a drop of water*, may, for present purposes, be taken to be a simple natural kind predicate, whereas the other predicate is a semantically compound expression that is necessarily

<sup>17</sup> For a brief discussion of what happens when these semantical presuppositions fail, see the section “When Semantic Presuppositions Fail,” in chapter 10 of *Beyond Rigidity*.

<sup>18</sup> One should also be careful about taking this demonstration of how the necessity of certain statements can be inferred from their truth to be a story about how we go about discovering the truth and necessity of statements of this sort in practice. A reasonable idealized model of this might include (i) investigation of the A-sample to determine that most of its members are instances of a unique kind k of the relevant type T, (ii) projection of the extension of A from this finding, (iii) similar determination of the kind B and the extension of B, and (iv) comparison of kinds and extensions of the two predicates, (v) determination of truth of the claim *All A's are B's*, and (vi) use of relevant elements of the above model to derive the necessity of this claim.

coextensive with the kind designated by that simple predicate.<sup>19</sup> Because this compound predicate is not semantically simple, (13) is not an instance of paradigm (11). Thus, even though (13) is an example that is now (after Kripke) widely regarded as both necessary and a posteriori, nothing we have said up to now shows that its necessity has anything to do with the non-descriptive semantics assigned to the simple natural kind predicate *is a drop of water* by Kripke's theory. But why should the complexity of the predicate on the right-hand side of (13) make a difference? Shouldn't we be able to apply the reasoning behind (11) directly to (13), despite the fact that the B-predicate is semantically complex? The restriction in (11) to cases in which the B-predicate is semantically simple seems inessential. Surely, it might be suggested, it is enough that B be a natural kind predicate. Isn't the predicate *is a drop of a substance molecules of which contain two hydrogen atoms and one oxygen atom* such a predicate?

At this point, before jumping to any conclusions, we would do well to compare (13) to (14) and (15).

14. For all x, x is a drop of water iff x is a drop of the substance instances of which fall from the sky in rain and fill the lakes and rivers.
15. For all x, x is a drop of water iff x is a drop of any one of the liquids that fall from the sky in rain and fill the lakes and rivers.

These examples are, we may assume, true but contingent. Thus the reasoning in (11) cannot apply to them. But why not? First consider (14). The predicate on the right-hand side of this example contains a singular definite description that non-rigidly designates a certain substance. With respect to the actual state of the world, that substance is water; with respect to another world-state, a different substance may be designated. Since the substance designated by this description with respect to an arbitrary world-state is the kind that determines the extension of the compound predicate at the world-state, different kinds determine the extension of the predicate with respect to different world-states. As a result, the reasoning behind paradigm (11) fails at step (ii) (and step (vii)). Hence (14) is not characterized as necessary.

<sup>19</sup> This involves a simplification. In reality, the predicate *is a drop of water* is compound, in virtue of individuating water by the drop. I ignore this here, since it does not affect our central point. For a discussion of this issue, including an account of the semantics of simple mass predicates like *is water*, see chapter 11 of *Beyond Rigidity*.

Next consider (15). In this case, the compound predicate on the right-hand side doesn't contain any singular definite description, or any singular term, that either rigidly or non-rigidly designates any substance. Although, in fact, there is only one kind of liquid that falls from the sky in rain and fills the lakes and rivers, this substance—water—is not semantically associated with the predicate in the way required by paradigm (11). Rather, the compound predicate expresses a property which, as it happens, is instantiated by all and only instances of the substance water in the world as it actually is, but which could have been instantiated by a different substance, or even by a motley collection of different liquid substances, mixtures, and the like in different possible circumstances. Thus, the reasoning behind paradigm (11) again fails at step (ii) (and (viii)), and (15) is not characterized as necessary.

What distinguishes (13) from (14) and (15)? Why are we inclined to think that the former is necessary, whereas the latter pair are not? The answer has to do with what we believe about substances. As Nathan Salmon pointed out many years ago, we believe that it is a feature of any genuine substance *S* that whatever its molecular structure turns out to be, all possible instances of *S* share that structure, and all possible instances of that structure are instances of *S*.<sup>20</sup> From this plus the truth of (13) and the semantics of 'water', it will follow that (13) is necessary. But what is a substance, and why do we believe molecular structure to be constitutive of it?

Here is a speculative proposal (which goes beyond anything explicitly indicated by Kripke). A substance is one type of natural kind. Natural kinds are abstract objects which have individual concrete entities as instances. Whatever these kinds are, they are individuated by their possible instances in the following (weak) sense: for all natural kinds *x* and *y*,  $x \neq y$  iff there is some possible state of the world *w* such that the class of things that are instances of *x* with respect to *w*  $\neq$  the class of things that are instances of *y* with respect to *w*. In other words, kinds with the same possible instances are identical. By contrast, properties—particularly those expressed by semantically compound predicates like *is an equiangular triangle* and *is an equilateral triangle*—may differ even if they are necessarily coextensive. Even though properties are more fine-grained than natural kinds, they may determine kinds. We will say that a property *P* determines a natural kind *k* iff for all possible states of the world *w*, the extension of *P* with respect to *w* is just the extension

<sup>20</sup> Nathan Salmon, *Reference and Essence* (Princeton, NJ: Princeton University Press, 1981).

of *k* with respect to *w*. (Note that properties may determine at most one natural kind, whereas a given natural kind may have many properties that determine it.) Substances are natural kinds determined by properties of a special type—namely, properties of an individual, or sample of matter, *x* that specify how *x* is made, or constituted, out of basic physical stuff.

With this in mind, consider the following idealized story about the introduction of the simple natural kind term *water*. Let us suppose that it is introduced with the intention that it is to be a substance term that applies to everything that shares the same physical constitution as the individual samples of water that led to its introduction.<sup>21</sup> On this view, when we introduce the term, we may know neither what physical constituents make up these samples, nor how they are combined. Nevertheless, we intend the term to apply (with respect to a possible world-state) to all and only those things that share the basic physical constitution (whatever it may turn out to be) exhibited (in the world as it actually is) by all, or nearly all, elements in our sample. We may put this by saying that when we introduce the term *water*, we, in effect, stipulate that it is to apply to instances of the substance—the unique physically constitutive kind—of which nearly all members of the sample are instances.<sup>22</sup>

On this idealized model, it may well be that no one knows much about the nature of this physically constitutive kind when the predicate *is a drop of water* is introduced. Consequently, it is no threat to the model that the predicate was introduced long before the development of chemical theories of molecular structure. When these theories were developed, we acquired new families of concepts to describe these structures, and we came to understand claims about the molecular structures of macroscopic objects as claims about how those objects are constituted out of basic physical constituents. In short, these theories

<sup>21</sup> For a discussion of a somewhat less idealized scenario concerning the introduction of *water*, which does not require those introducing the term to have explicit intentions about physical constitution, see the section entitled “The Role of Intention in Determining the Reference of Natural Kind Terms” in chapter 10 of *Beyond Rigidity*.

<sup>22</sup> Strictly speaking, this explanation applies to one, expansive, sense of the term *water*. This term also has a restricted sense in which it applies to all and only instances of the substance water that are in liquid form. In this sense, it contrasts with *ice*, which applies to all and only instances of the substance water in frozen form. *Water*, when used with its expansive sense, is a simple substance term; when used with its restricted sense, it, like *ice*, is partially descriptive. In order to explain these distinctions properly, one must deal with the use of the terms *water* and *ice* as simple mass predicates. This is done in chapter 11 of *Beyond Rigidity*, where a more complete theory is given, and various puzzles and problems are dealt with.

brought with them a class of formulas expressing properties which, by virtue of specifying different possible molecular structures, were understood as determining different possible physically constitutive kinds or substances. After empirical investigation, certain formulas—e.g., *molecules of  $x$  contain two hydrogen atoms and one oxygen atom*—were taken to be candidates for giving the basic physical constitution of standard water samples. In this way, the predicates *is a drop of water* and *is a drop of a substance molecules of which contain two hydrogen atoms and one oxygen atom* came to be understood as designating physically constitutive kinds, and hence as designating kinds of the same type. This had the consequence that when it was discovered empirically that water samples have the molecular structure  $H_2O$ , this was sufficient to guarantee that the two predicates designated the same kind—in the case of *water*, the kind is designated by virtue of being the very meaning of the term, whereas in the case of the compound predicate, the kind is designated by virtue of being determined by the property it expresses. Since this kind may be seen as determining the extensions of both predicates with respect to all possible world-states, it follows, without need of further assumptions, that (13) is a necessary truth.

The same sort of explanation might be given for related examples, such as (16).

16. Chunks of gold are chunks of an element the atomic number of which is 79.

What about other natural kind predicates that are not substance terms? One familiar sort of case involves what might be called *explanatory kinds*, where these are thought of as kinds determined by properties the possession of which causally explains certain observed characteristics. Heat and molecular motion, in Kripke's discussion, seem to fall into this category. To see how this is supposed to go, imagine that the abstract singular term *heat* is introduced by an ostensive definition stipulating that it is to designate the unique kind of physical state or process present in a certain class of samples that causes certain effects, including, among other things, certain reactions in us. Imagine the further stipulation that the related predicate *is hotter than* is to apply to pairs in which the relevant physical state or process in one object is more pronounced than it is in the other object. When the kinetic theory is formulated as a hypothesis to be tested, it will be antecedently understood that a relational predicate characterizing different levels of mean molecular kinetic energy is a candidate for specifying the natural

kind designated by *is hotter than*. Once it is discovered that certain facts about kinetic energy causally explain certain effects, it will follow that the kind designated by *is hotter than* just is the kind determined by the relation (i.e., the 2-place property) having-a-higher-mean-molecular-kinetic-energy-than, and Kripke's example (12) will be characterized not only as true, but also as necessary.

This explanation of the necessity of (12) is slightly different from the explanation of the necessity of (13), which was built on paradigm (11). When an explanatory kind is involved, the explanation runs roughly as follows: (i) A simple natural kind term E is introduced with the stipulation that it applies (with respect to any possible state of the world) to all and only instances of the kind determined by those properties possession of which (causally) explains certain phenomena in the world as it actually is. (ii) It is then discovered scientifically that the property of being such and such (causally) explains the phenomena. (iii) From this it follows that the kind directly designated by the simple predicate E is the kind determined by the property of being such and such. (iv) This is sufficient to establish the necessity of the claim that says that an individual (or a pair) "is E" iff it is (or they are) such and such.

### **The Epistemic Status of Theoretical Identities Involving Natural Kind Predicates**

So far we have constructed an account, using Kripke's analysis of natural kind terms, to explain the necessity of certain theoretical identification sentences containing natural kind predicates. It remains to determine whether the necessary truths expressed by these sentences are knowable a priori, or only a posteriori. Let us begin with (13). In explaining why it is necessary, we noted that the two predicates it contains—*is a drop of water* and *is a drop of a substance molecules of which contain two hydrogen atoms and one oxygen atom*—designate the same natural kind. However, we also noted in passing that these predicates have different meanings. Whereas the meaning of *water* is simply the natural kind that it designates, the meaning of *substance molecules of which contain two hydrogen atoms and one oxygen atom* is a (complex) property that determines that kind. Since the two predicates mean different things, (13) is not synonymous with, and does not express the same proposition as, the trivial identity (13a).

13a. For all x, x is a drop of water iff x is a drop of water.

Thus, the fact that the necessary proposition expressed by (13a) is knowable apriori does not show that the necessary proposition expressed by (13) is knowable apriori. That is all to the good.

If, in addition, the kind water is such that one cannot know of it, except by empirical investigation, that for all  $x$ ,  $x$  is an instance of it iff  $x$  is made up of molecules containing two hydrogen atoms and one oxygen atom, then the proposition expressed by (13) is **not** knowable apriori, but rather is a genuine example of the necessary aposteriori. This supposition is not unreasonable. After all, we don't know the kind water by mere reflection, or by virtue of being given a verbal definition of any sort. Rather, our acquaintance with it is as something shared by a variety of concrete particulars which we use to identify it. This being so, one's knowledge that drops of water are drops of  $H_2O$  must be grounded, either in one's own knowledge that particular water samples have the molecular structure  $H_2O$ , or in the knowledge of others, passed on to one, which they are able to ground in their knowledge of particular water samples. This knowledge can only be aposteriori.

An important part of this explanation is the account of why the two predicates in (13) do not mean the same thing. In giving this account, I have moved well beyond doctrines that Kripke explicitly endorses. In effect, I have suggested that we adopt a position that may justly be called *Extended Millianism*. Simple Millianism, which was inspired in its modern form by Kripke, but not endorsed by him, is the view that the meaning of a proper name is its referent. On this view, coreferential names mean the same thing, and true identity sentences involving them semantically express propositions that are necessary and apriori, rather than necessary and aposteriori (though such sentences may nevertheless be used in different contexts to assert or convey various aposteriori truths as well).<sup>23</sup> Extended Millianism holds that the meaning of a simple natural kind predicate is the natural kind it designates; as a result, simple predicates that designate the same natural kind—such as *groundhog* and *woodchuck*—mean the same thing. This has the immediate consequence that the theoretical identity sentence

17a. For all  $x$ ,  $x$  is a woodchuck iff  $x$  is a groundhog.

which fits paradigm (11), and hence is a necessary truth, semantically expresses the same proposition as

<sup>23</sup> This view is defended by Nathan Salmon in *Frege's Puzzle* (Cambridge, MA: MIT Press, 1986), and by me in chapters 1–8 of *Beyond Rigidity*.

17b. For all  $x$ ,  $x$  is a woodchuck iff  $x$  is a woodchuck.

and hence is knowable apriori (though it too can be used to assert or convey various aposteriori truths in different contexts).

Things change when we consider compound expressions. Just as in the case of noun phrases, singular definite descriptions have meanings that are not identical with the objects they denote, but rather are properties in virtue of which they denote what they do, so a semantically compound natural kind predicate, such as *is a drop of a substance molecules of which contain two hydrogen atoms and one oxygen atom*, has a meaning that is not identical with the natural kind it designates, but rather is a property that determines that kind. On this picture, simple natural kind predicates are analogous to proper names, and compound descriptive predicates are analogous to singular definite descriptions. Because of this, it is possible for some sentences that have the logical form  $\forall x (Ax \leftrightarrow Bx)$  to be both necessary, if true, and knowable only aposteriori—where one of the predicates is a simple natural kind predicate, and the other is a compound descriptive phrase.

### A Final Challenge to the Necessary Aposteriori

Before leaving *Naming and Necessity*, there is one final discussion, not far from the end of the work, that merits our close attention. After summing up his treatment of natural kind terms, and illustrating their role in generating surprising and significant examples of the necessary aposteriori, Kripke takes up a fundamental challenge to his conception of this class of truths. Up until this point in the lectures, when describing what he takes to be instances of the necessary aposteriori, he emphasizes that although they are necessary, and hence true with respect to every possible state of the world, nonetheless, for all we knew prior to empirically discovering their truth, “they could have turned out otherwise.” Realizing that this may sound puzzling and problematic, he gives voice in lecture 3 to the following objection.

Theoretical identities, according to the conception I advocate, are generally identities involving two rigid designators, and therefore are examples of the necessary *a posteriori*. Now in spite of the arguments I gave before for the distinction between necessary and *a priori* truth, the notion of *a posteriori* necessary truth may still be somewhat puzzling. Someone may well be inclined to argue as



follows: ‘You have admitted that heat might have turned out not to have been molecular motion, and that gold might have turned out not to have been the element with the atomic number 79. For that matter, you also have acknowledged that . . . this table might have turned out to be made from ice from water from the Thames. I gather that Hesperus might have turned out not to be Phosphorus. What then can you mean when you say that such eventualities are impossible? If Hesperus might have *turned out* not to be Phosphorus, then Hesperus might not have *been* Phosphorus. And similarly for the other cases: if the world could have *turned out* otherwise, it could have *been* otherwise.’<sup>24</sup>

The problem here starts out being about theoretical identity sentences involving natural kind predicates, but quickly expands to cover all instances of the necessary aposteriori. Let *p* be such an instance. Since *p* is aposteriori, its falsity must be conceivable, and we need empirical evidence to rule out that possibility. Without such evidence **it could turn out that**, for all we know, *p* is false. But, the objector maintains, if *p* is necessary, there are no possibilities to be ruled out in which *p* is false, since no matter what possible state the world could be in, it is a state in which *p* is true. Thus, if *p* really is necessary, we don’t require empirical evidence to know *p* after all, and if *p* really is aposteriori, then *p* isn’t necessary; either way, the necessary aposteriori is an illusion.

Kripke begins his reply to this objection with the following passage.

The objector is correct when he argues that if I hold that this table could not have been made of ice, then I must also hold that it could not have turned out to be made of ice; *it could have turned out that P* entails that *P* could have been the case. What, then, does the intuition that the table might have turned out to have been made of ice or of anything else, that it might even have turned out not to be made of molecules, amount to? I think that it means simply that there might have been *a table* looking and feeling just like this one and placed in this very position in the room, which was in fact made of ice. In other words, I (or some conscious being) could have been *qualitatively in the same epistemic situation* that in fact obtains, I could have the same sensory experience that I in fact have, about *a table* which was made of ice.<sup>25</sup>

<sup>24</sup> *Naming and Necessity*, pp. 140–41. Here and throughout, italic emphasis is Kripke’s.

<sup>25</sup> *Ibid.*, pp. 141–42.

In discussing what Kripke says, it is helpful to imagine the following scenario: A table has been brought into my office; I have examined it and determined that it is made out of wood, and not ice. I point to the table and truly say that I know that it is not made out of ice. I know that it is not made out of ice because I have used empirical observation and investigation to rule out what otherwise would have been an epistemologically relevant possibility. Prior to that investigation, **it could have turned out**, for all I knew, that the table was made of ice. What does Kripke tell us about this? He tells us that the intuition that **it could have turned out** that this table was made of ice is nothing more than the judgment that it is genuinely possible for me, or some other agent, to be in a situation qualitatively identical to the one I am in now, and be pointing at a table that is made out of ice.

He generalizes this point in the next paragraph.

The general answer to the objector can be stated, then, as follows: Any necessary truth, whether *a priori* or *a posteriori*, could not have turned out otherwise. In the case of some necessary *a posteriori* truths, however, we can say that under appropriate qualitatively identical evidential situations, **an appropriate corresponding qualitative statement** might have been false. The loose and inaccurate statement that gold might have turned out to be a compound should be replaced (roughly) by the statement that it is logically possible that there should have been a compound with all the properties originally known to hold of gold. The inaccurate statement that Hesperus might have turned out not to be Phosphorus should be replaced by the true contingency mentioned earlier in these lectures: two distinct bodies might have occupied, in the morning and the evening, respectively, the very positions actually occupied by Hesperus-Phosphorus-Venus.<sup>26</sup>

This paragraph, together with the one just preceding it, marks the beginning of what, in my opinion, is potentially the most misleading, and even disastrous, passage in *Naming and Necessity*.

Two main issues are addressed: the necessity of certain propositions, and the fact that they are knowable only *aposteriori*. Regarding the former, Kripke makes the following points:

(i) that there is a natural and correct way of understanding the locution *It could have turned out that ~S* in which it entails *It is not necessary that S*,

<sup>26</sup> Ibid., 142–43, my boldface emphasis.

(ii) that when the locution is understood in this way, his previous remarks—to the effect that empirical evidence is needed to rule out possible circumstances responsible for the truth of *It could have turned out that*  $\sim S$  in cases in which  $S$  is an example of the necessary aposteriori—were strictly speaking, inaccurate, but

(iii) that even when *it could have turned out that* is understood as in (i), and  $S$  is a genuine example of the necessary aposteriori, there often are different, descriptive propositions which are both contingent and knowable only aposteriori, and which are also easily confused with the proposition expressed by  $S$ ; these are the propositions the negations of which could genuinely have turned out to be true.

These points provide the basis for Kripke's answer to the objector. He maintains that when the objector protests that his examples of the necessary aposteriori can't be necessary if they are really aposteriori, he is **confusing** the propositions expressed by the examples with other, related propositions that, although knowable only aposteriori, are really contingent. In the case of my claim that this table is made out of ice, the objector is confusing it with the proposition that the table in front of me that appears so and so is made out of ice, or some closely related proposition. In the case of Kripke's claim that the statement that Hesperus is Phosphorus is both necessary and aposteriori, the objector is confusing it with the statement that the heavenly body that appears in the evening sky (at certain places and times) is the heavenly body that appears in the morning sky (at other places and times), and so on for the other examples.

Although this response is unobjectionable, as far as it goes, it does not go far enough. Two difficulties stand out. First, there is a natural way of understanding the worry raised by the objector in which it is general, and not dependent on any particular example about which one might be confused. If, as seems plausible, the function of the empirical evidence needed to justify aposteriori knowledge is to rule out possible circumstances in which the proposition known is false, then it would seem to follow that there simply can't be any necessary truth for which empirical evidence is necessary in order for one to know it.<sup>27</sup>

<sup>27</sup> A general objection of this sort can easily be extracted from the definition of knowledge proposed by David Lewis in "Elusive Knowledge," *Australasian Journal of Philosophy* 74 (1996), reprinted in his *Papers in Metaphysics and Epistemology* (Cambridge: Cambridge University Press, 1999). In addition, that paper contains another basis that Lewis and some other philosophers have used to object to the idea that there are necessary truths that are knowable

Since this problem naturally springs to mind when considering the position of the objector, one would expect Kripke to address it and provide a solution. That he does not do so here is striking. In focusing so much on the necessity of his examples, he appears to neglect the question of how these necessary truths are known, and in particular, the role of empirical evidence in coming to know them.

Second, to the extent that he does suggest an answer to this question, it is puzzling and inadequate. In considering the Hesperus/Phosphorus example, he directs us to his earlier discussion at the end of lecture 2. However, as we saw in chapter 15, the explanation he gives there of the aposteriori character of this example is seriously flawed. As I argued earlier, it appears that his argument that the proposition that Hesperus is Phosphorus is knowable only aposteriori implicitly relies on some version of what I called *principles of strong disquotation and justification*. Details aside, the idea expressed by these principles is that if I am a competent speaker who understands S, then I satisfy the formula *x is justified in believing that S* on the basis of evidence e iff I accept S, believe it to be true, and am justified in so doing on the basis of e. In the Hesperus/Phosphorus case, if I understand the sentence *Hesperus is Phosphorus*, while associating the two names with the descriptions *the heavenly body seen in the evening sky (at certain places and times)*, and *the heavenly body seen in the morning sky (at other places and times)*, then I will justifiably accept and believe the sentence *Hesperus is Phosphorus* to be true only if I justifiably believe that the heavenly body seen in the evening sky (at certain places and times) is the heavenly body seen in the morning sky (at other places and times). Since my justification for this descriptive belief can only be empirical, my justification for accepting the sentence *Hesperus is Phosphorus* must also be empirical. Strong disquotation and justification will then tell us that my belief that Hesperus is Phosphorus is one that is justified empirically, and hence that my knowledge of this proposition is aposteriori, rather than apriori. Apparently believing that that this result generalizes to other agents, times, and sentences expressing the proposition that Hesperus is Phosphorus, Kripke arrives at the conclusion that this proposition can be known only aposteriori.

---

only aposteriori—namely, the identification of propositions with sets of metaphysically possible worlds-states. For arguments against any such identification, see my “Direct Reference, Propositional Attitudes, and Semantic Content,” originally published in *Philosophical Topics* 15 (1987):47–87; reprinted in N. Salmon and S. Soames, eds., *Propositions and Attitudes* (Oxford: Oxford University Press, 1988).

As I indicated in chapter 15, the problem with this argument is that, depending on the specific versions in question, principles of strong disquotation and justification are false, or at least highly questionable. Thus, Kripke's explanation of the alleged aposterioricity of the necessary truth that Hesperus is Phosphorus cannot be accepted. That was bad enough at the end of lecture 2, when one putative example of the necessary aposteriori was in question. Here in lecture 3, when he is responding to a general objection to all examples of the necessary aposteriori, it is disastrous. Here, the suggestion implicit in his discussion is that the earlier explanation of the alleged aposterioricity of our knowledge that Hesperus is Phosphorus is to be generalized to all cases of the necessary aposteriori.

To see how this might go, consider the example of the table that has been brought into my office. In pointing to the table and saying *This table is not made out of ice*, I am expressing a proposition that is a necessary truth—since the particular table is one that could not have existed if it were made out of ice. Nevertheless, in the circumstances in which I find myself, I would not accept, and would not be justified in accepting, the sentence *This table is not made out of ice* (where *this* refers to the table in question) unless I also believed, and was justified in believing, the general, descriptive proposition that the (or a) table that has just been brought into my office, which appears so and so, is not made out of ice. This proposition is, of course, contingent rather than necessary, and hence not to be confused with the proposition expressed by the indexical sentence I uttered. However, it is also one that I am justified in believing only on the basis of empirical evidence. Since this evidence is included in the evidence on which I base my utterance, my evidence for accepting the sentence uttered, and believing it to be true, must also be empirical. Appealing to principles of strong disquotation and justification, one gets the result that although it is a necessary truth that this table is not made out of ice, and although I know that this table is not made out of ice, my knowledge is based on empirical evidence, and hence is aposteriori.<sup>28</sup> Generalizing to other agents, times, and alternative ways of expressing the same proposition, one might naturally conclude that the proposition that this table is

<sup>28</sup> For this one needs context-relativized principles of strong disquotation, and strong disquotation and justification. Kripke, of course, does not explicitly formulate any such principles. For an indication of one way to formulate a context-relativized version of strong disquotation, and a discussion of problems with such principles, see pp. 12–13 of *Beyond Rigidity*.

made out of ice is one that is both necessary and knowable only aposteriori. On this view, there may be no single contingent, aposteriori, and descriptive proposition  $q$  which is such that for every agent  $a$ , in order for  $a$  to know that this table is not made out of ice,  $a$  must know  $q$ ; nevertheless, in order for any agent  $a$  to know that this table is not made out of ice,  $a$  must know some contingent, aposteriori, and descriptive proposition  $q$  that  $a$  associates with this proposition. This, it might be thought, is sufficient to guarantee that knowledge that this table is not made out of ice can only be aposteriori.

One piece of evidence that Kripke intended to convey some such view as this is that he seemed to see his response to the objector as generalizing his earlier treatment of the Hesperus/Phosphorus example to all of his other examples of the necessary aposteriori. This is suggested in the paragraph immediately following the one last cited.

I have not given any general paradigm for the appropriate corresponding qualitative contingent statement. Since we are concerned with how things might have turned out otherwise, our general paradigm is to redescribe both our prior evidence and the statement qualitatively and claim that they are only contingently related. In the case of identities, using two rigid designators, such as the Hesperus-Phosphorus case above, there is a simpler paradigm which is often usable to at least approximately the same effect. Let ' $R_1$ ' and ' $R_2$ ' be the two rigid designators that flank the identity sign. Then ' $R_1 = R_2$ ' is necessary if true. The references of ' $R_1$ ' and ' $R_2$ ', respectively, may well be fixed by nonrigid designators ' $D_1$ ' and ' $D_2$ ', in the Hesperus and Phosphorus cases these have the form 'the heavenly body in such-and-such position in the sky in the evening (morning)'. Then although ' $R_1 = R_2$ ' is necessary, ' $D_1 = D_2$ ' may well be contingent, and this is often what leads to the erroneous view that ' $R_1 = R_2$ ' might have turned out otherwise.<sup>29</sup>

In the case of the Hesperus/Phosphorus example, Kripke suggests that the related contingent proposition with which the objector confuses the necessary truth that Hesperus is Phosphorus is also a proposition that the agent must know in order to be counted as knowing that Hesperus is Phosphorus. Here, he seems to suggest that the same point can be made for all of his other examples of the necessary aposteriori as well.

This, I believe, is Kripke's final answer to the objector in lecture 3.

<sup>29</sup> *Naming and Necessity*, pp. 143–44.

Unfortunately, it is highly problematic. Although a number of lesser worries might be raised, the most serious problem is that Kripke's answer extends the inadequacies found in chapter 15 with his account of the Hesperus/Phosphorus example to all of his examples of the necessary aposteriori—thereby threatening to undermine one of the most important general philosophical theses in *Naming and Necessity*. Fortunately, this result can be avoided. Although the answer that Kripke actually gives to the objector is problematic, there is another, much more plausible answer he could have given that grows naturally out of his general philosophical position. Recall the objector's point. If *p* is knowable only aposteriori, then empirical evidence is needed to rule out certain possible circumstances in which *p* is false. However, if *p* is necessary, then there are no such circumstances to rule out, since *p* is true with respect to every possible state that the world could be in. Thus, no proposition can be both necessary and knowable only aposteriori. Faced with this argument, there are three main responses that could be offered by a defender of the necessary aposteriori: (i) he could reject the idea that when empirical evidence is required for knowledge, its function is to rule out possibilities; (ii) he could reject the idea that all conceptual, or epistemic, possibilities are genuine, metaphysical possibilities—i.e., he could reject the claim that every way that, for all we know, the world might be is a way that the world genuinely could be; or, (iii) he could maintain that examples of the necessary aposteriori are those in which even though a sentence *S* expresses a necessary truth *p*, what we standardly call *knowing that S* always requires knowing some contingent, aposteriori proposition *q* that is related to *p* in a certain way. As I have argued, Kripke seems to have chosen (iii).<sup>30</sup> However, given his other views, his most natural response should have been to choose (ii)—to reject the idea that all epistemic possibilities are genuine metaphysical possibilities. He is committed to this anyway; and since it is all he needs to block the objection, no other response is necessary.<sup>31</sup>

<sup>30</sup> The fact that Kripke's discussion can, without much strain, be read as eventuating in (iii) has lent some support to those who maintain that no single proposition can be both necessary and knowable only aposteriori. See, for example, Robert Stalnaker, "Assertion," *Syntax and Semantics* 9 (1978):315–32, reprinted in his *Context and Content* (New York: Oxford University Press, 1999); David Lewis, "Elusive Knowledge," *Papers in Metaphysics and Epistemology* (Cambridge: Cambridge University Press, 1999); and Frank Jackson, *From Metaphysics to Ethics* (New York: Oxford University Press, 1998). Criticism of option (iii), in its various manifestations, can be found in my *Reference and Description: The Two-Dimensionalist Attempt to Revive Descriptivism* (in preparation), as well as in my "Saul Kripke, the Necessary Aposteriori, and the Two-Dimensionalist Heresy."

<sup>31</sup> I am indebted to Ali Kazmi for very helpful discussions of this section.

## A Proper Response to the Challenge: Essential Properties, Impossible World-States, and the Necessary Aposteriori

As I indicated in chapter 15, Kripke's views about the necessary aposteriori are connected to his views about essential properties.<sup>32</sup> He argues that we know *apriori* that various properties and relations—such as non-identity, being human, not being made out of ice, and being made out of molecules—are essential to anything that has them.<sup>33</sup> Thus, we know *apriori* that if things have these properties and relations, then they have them necessarily. This means that various propositions which predicate these properties and relations of objects are such that we know *apriori* that **if** they are true, **then** they are necessarily true. Still, finding out that they are true requires empirical investigation. If this is right, then sometimes in order to find out whether certain things are true with respect to all possible states of the world, and other things are true with respect to no possible states of the world, we first must find out what is true with respect to the actual state of the world. Sometimes in order to find out what could or could not be, one first must find out what is. This will seem problematic only if you have restricted the ways things could **conceivably** be to the ways things could **really** be—i.e., only if you have restricted epistemic possibility to metaphysical possibility. Although the passage in lecture 3 of *Naming and Necessity* that we have been talking about may seem to show Kripke ignoring or backsliding on this point, it doesn't, in my opinion, negate the central lesson of his work that one must sharply distinguish these two kinds of possibility.<sup>34</sup> Once this is

<sup>32</sup> See in particular in his essay "Identity and Necessity."

<sup>33</sup> There are also properties that are essential to some, but not all, things that have them. These too give rise to instances of the necessary aposteriori that can be treated in roughly the manner suggested in the text. In the interest of simplicity, I put these to one side in the text.

<sup>34</sup> Kripke's footnote 72, toward the end of the passage under discussion, shows that even there he was aware of the importance of the distinction between epistemic and metaphysical possibility. He says, referring to some of his remarks in the passage: "Some of the statements I myself make above may be loose and inaccurate in this sense. If I say, 'Gold *might* turn out not to be an element,' I speak correctly; 'might' here is *epistemic* and expresses the fact that the evidence does not justify *a priori* (Cartesian) certainty that gold is an element. I am also strictly correct when I say that the elementhood of gold was discovered *a posteriori*. If I say, 'Gold *might have* turned out not to be an element,' I seem to mean this metaphysically and my statement is subject to the correction noted in the text." To understand the relation between this footnote and the text, it is important to remember that the footnotes were added to the lectures by Kripke a considerable time after they were given and a written transcript had been produced. I believe that when writing the footnote, he noticed that his discussion in the text had neglected the important and relevant distinction between epistemic and metaphysical necessity, and he wished—without changing the text—to call the reader's attention to his commitment to this distinction.



done, and both rigid designation and the existence of nontrivial essential properties of objects is accepted, the necessary *a posteriori* follows unproblematically.

This point is connected to the initial idealized conception of inquiry discussed in chapter 15, which is presupposed by Kripke's hypothetical objector in lecture 3. On that conception, ignorance is a matter of lacking information about which, of certain different possible states that the world could be in, it is actually in; and complete ignorance is a condition in which one doesn't know which, of all the possible states that the world could be in, it is actually in. According to this conception, when an agent is in this condition, (i) all metaphysically possible states of the world are epistemically possible—i.e., every way that the world could possibly be is a way that, for all the agent knows, it actually is, and (ii) every epistemic possibility is a metaphysical possibility—i.e., every way that, for all the agent knows, the world might be is a way that the world genuinely could be. Inquiry is the process of escaping from ignorance. By investigating the world, the agent learns contingent truths that distinguish the way the world actually is from other ways it might possibly be, but isn't. Each time the agent learns one of these truths, he narrows down the class of possibilities compatible with what he knows, and within which he locates the way the world actually is. Thus, acquiring information is equated with narrowing down the range of genuinely possible world-states that are compatible with what one knows. On this picture, a proposition *q* provides evidence supporting the truth of a proposition *p* by ruling out certain possible ways in which *p* might fail to be true. This has the immediate consequence that the idea of necessary truths which are knowable only *a posteriori* becomes problematic. It is problematic because to say that *p* is knowable only *a posteriori* is to say that one can have the justification required to know *p* only if one has empirical evidence supporting its truth. Since *p* is necessary, however, there are no possible world-states with respect to which it is untrue; hence there can be no empirical evidence supporting *p*. Therefore, there is no such thing as the necessary *a posteriori*.

The objector in lecture 3 is naturally read as having something like this in mind. Hence, it is strange that Kripke didn't challenge the presuppositions underlying that conception of inquiry. Instead, he ended up appearing to suggest that what makes it possible for a necessary truth to be knowable only *a posteriori* is that, in the end, knowledge of it always involves knowledge of something else—something contingent

which requires empirical support. However, there was no need for this detour. Here, it is helpful to remember that for Kripke, possible states of the world are not alternate concrete universes, but abstract objects. As I pointed out in chapter 15, they can be thought of as maximally complete ways the real concrete universe could have been—maximally complete properties that the universe could have instantiated. Thinking of them in this way suggests an obvious generalization. Just as there are properties that certain objects could possibly have and other properties that they couldn't possibly have, so there are certain maximally complete properties that the universe could have had—possible states of the world—and other maximally complete properties that the universe could not have had—impossible states of the world. Just as some of the properties that objects couldn't have had are properties that one can conceive them as having, so some maximally complete properties that the universe could not have had (some impossible states of the world) are properties one could conceive it as having. Given this, one can explain the informativeness of certain necessary truths as resulting from the fact that learning them allows one to rule out certain impossible, but nevertheless conceivable, states of the world. Moreover, one can explain the function played by empirical evidence in providing the justification needed for knowledge of necessary *a posteriori* truths; empirical evidence is required to rule out certain impossible, but nevertheless conceivable and epistemologically relevant, world-states with respect to which the necessary propositions are false.<sup>35</sup> Thus, by expanding the range of epistemically conceivable states of the world to include some that are metaphysically impossible, one can modify the original conception of inquiry so as to accommodate Kripkean examples of the necessary *a posteriori*. Whether or not this is the final word on the subject is another question. As indicated at the end of chapter 15, some further modification even of this expanded conception of inquiry may well be needed. However, the central idea that not all epistemic possibilities are metaphysical possibilities seems both to be solid and to provide the key to responding to Kripke's hypothetical objector to the necessary *a posteriori*.

Once one dispels the air of mystery surrounding this category of truth, the motivation for Kripke's problematic response to the objector in lecture 3 falls away, and the full force of his discovery of the necessary

<sup>35</sup> As in chapter 15, I here assume that names (unlike definite descriptions such as *the stuff out of which this table, if it exists, is constituted*) rigidly designate the same thing with respect to all world-states, possible or not.

aposteriori can be appreciated. One of the great philosophical achievements of the twentieth century, it has transformed the philosophical landscape, recalibrated our sense of what is possible, and reshaped our understanding of our own philosophical past. No single insight has been more important than this in gaining the perspective needed to understand and critically evaluate the philosophical tradition stretching from Moore, Russell, and Wittgenstein, through logical positivism and the ordinary language school, to Quine, Davidson, and Kripke himself. Without Kripke's discovery, the history told in these pages would have been very different; indeed, **these volumes** would scarcely have been possible.

## SUGGESTED FURTHER READING FOR PART SEVEN

### Main Primary Sources Discussed

Kripke, Saul. *Naming and Necessity*. Cambridge, MA: Harvard, 1980; originally published in D. Davidson and G. Harman, eds., *Semantics of Natural Languages* (Dordrecht: Reidel 1972), pp. 253–355.

### Additional Primary Sources

Kripke, Saul. “Identity and Necessity.” In Milton Munitz, ed., *Identity and Individuation* (New York: NYU Press, 1971).  
———. “A Puzzle about Belief.” In A. Margalit, ed., *Meaning and Use*, (Dordrecht: Reidel, 1979); reprinted in N. Salmon and S. Soames, eds., *Propositions and Attitudes* (Oxford: Oxford University Press, 1988).

### Additional Recommended Reading

Burgess, John. “How Not to Write History of Philosophy.” In *The New Theory of Reference* (Dordrecht, Boston, and London: Kluwer, 1998).  
———. “Marcus, Kripke, and Names.” In Paul W. Humphreys and James H. Fetzer, eds., *The New Theory of Reference*.  
———. “Quinus ab omni naevo vindicatus.” In Ali A. Kazmi, ed., *Meaning and Reference* (Calgary: University of Calgary Press, 1998).  
Chalmers, David. “On Sense and Intension.” *Philosophical Perspectives* 16, *Language and Mind* (Oxford: Blackwell, 2002), 135–82.  
Evans, Gareth. “The Causal Theory of Names.” *Proceedings of the Aristotelian Society*, supplementary vol. 47, 1973, 187–208; reprinted in his *Collected Papers* (Oxford: Oxford University Press, 1985).  
———. “Reference and Contingency.” In his *Collected Papers* (Oxford: Oxford University Press, 1985).  
Harman, Gil. “How to Use Propositions.” *American Philosophical Quarterly* 14 (1977):173–176.

- Jackson, Frank. *From Metaphysics to Ethics*. Oxford: Oxford University Press, 1998.
- Jeshion, Robin. "Acquaintanceless De Re Belief." In M. O'Rourke et al., eds., *Truth and Meaning: Investigations in Philosophical Semantics* (New York: Seven Bridges Press, 2001).
- Kaplan, David. "Demonstratives." In Joseph Almog, John Perry, and Howard Wettstein, eds., *Themes from Kaplan* (New York and Oxford: Oxford University Press, 1989).
- . "On The Logic of Demonstratives." *Journal of Philosophical Logic* 8 (1979):81–98.
- . "Opacity." In E. Hahn and P. A. Schilpp, eds., *The Philosophy of W. V. Quine* (La Salle, IL: Open Court, 1986).
- Koslicki, Kathrin. "The Semantics of Mass Predicates." *Noûs* 38 (1999): 46–91.
- Larson, Richard, and Peter Ludlow. "Interpreted Logical Forms." *Synthese* 95 (1993):305–56; reprinted in Peter Ludlow, ed., *Readings in the Philosophy of Language* (Cambridge, MA: MIT Press, 1997).
- Lewis, David. "Elusive Knowledge." In Lewis, *Papers in Metaphysics and Epistemology* (Cambridge: Cambridge University Press, 1999); originally published in *The Australasian Journal of Philosophy* 74 (1996).
- . "Naming the Colours." In his *Papers in Metaphysics and Epistemology*, (Cambridge: Cambridge University Press, 1999); originally published in *The Australasian Journal of Philosophy* 75 (1997).
- . *On the Plurality of Worlds*. Oxford: Blackwell, 1986.
- McKeown-Green, Jonathan. *The Primacy of Public Language*. Unpublished Princeton University Ph.D. dissertation, 2002.
- Neale, Stephen. "On a Milestone of Empiricism." In Alex Orenstein and Petr Kotatko, eds., *Knowledge, Language and Logic* (Dordrecht, Boston, and London: Kluwer, 2000).
- Putnam, Hilary. "Explanation and Reference." In G. Pearce and P. Maynard, eds., *Conceptual Change* (Dordrecht: Reidel, 1973); reprinted in Hilary Putnam, *Philosophical Papers*, vols. 1 and 2 (Cambridge: Cambridge University Press, 1975).
- . "Is Semantics Possible?" In H. Kiefer and M. Munitz, eds., *Language, Belief and Metaphysics* (Albany: State University of New York Press, 1970); reprinted in Putnam, *Philosophical Papers*.
- . "Meaning and Reference." *Journal of Philosophy* 70 (1973): 699–711; reprinted in Putnam's *Philosophical Papers*.
- . "The Meaning of 'Meaning'." In K. Gunderson, ed., *Lan-*

- guage, Mind, and Knowledge*, Minnesota Studies in the Philosophy of Science, no. 7 (Minneapolis: University of Minnesota Press, 1975); reprinted in Putnam, *Philosophical Papers*.
- Quine, W. V. "Notes on Existence and Necessity." *Journal of Philosophy* 40 (1943):113–27.
- . "The Problem of Interpreting Modal Logic." *Journal of Symbolic Logic* 12 (1947):43–48.
- . "Reference and Modality." In *From a Logical Point of View* (Cambridge, MA: Harvard University Press, 1953, 1961, 1980).
- . "Three Grades of Modal Involvement." In Quine, *The Ways of Paradox* (New York: Random House, 1966); originally published in 1953.
- Richard, Mark. *Propositional Attitudes*. Cambridge: Cambridge University Press, 1990.
- Rieber, Stephen. "Understanding Synonyms without Knowing That They Are Synonymous." *Analysis* 52 (1992):224–28.
- Salmon, Nathan. *Frege's Puzzle*. Cambridge, MA: MIT Press, 1986.
- . "How to Measure the Standard Metre." *Proceedings of the Aristotelian Society*, 88 (1987–88):193–217.
- . "A Millian Heir Rejects the Wages of Sinn." In C. A. Anderson and J. Owens, eds., *Propositional Attitudes: The Role of Content in Logic, Language, and Mind* (Stanford, CA: CSLI, 1990), 215–47.
- . "On the Logic of What Might Have Been." *Philosophical Review* 98 (1989):3–34.
- . *Reference and Essence*. Princeton, NJ: Princeton University Press, 1981.
- Soames, Scott. *Beyond Rigidity*. New York: Oxford University Press, 2002.
- . "Direct Reference, Propositional Attitudes, and Semantic Content," originally published in *Philosophical Topics*, 15, 1987, 47–87; reprinted in N. Salmon and S. Soames, eds., *Propositions and Attitudes* (Oxford University Press), 1988.
- . "Donnellan's Referential/Attributive Distinction." *Philosophical Studies* 73 (1994):149–68.
- . "More Revisionism about Reference." In *The New Theory of Reference* (Dordrecht, Boston, and London: Kluwer, 1998).
- . "Naming and Asserting." In A. Szabo, ed., *Semantics and Pragmatics*, (Oxford University Press, forthcoming).
- . *Reference and Description: The Two-Dimensionalist Attempt to Revive Descriptivism*. In preparation.

- . “Review of Gareth Evans, *Collected Papers*.” *Journal of Philosophy* 86 (1989).
- . “Revisionism about Reference.” *Synthese* 104 (1995):191–216; reprinted in *The New Theory of Reference* (Dordrecht, Boston, and London: Kluwer, 1998).
- . “Saul Kripke, the Necessary Aposteriori, and the Two-Dimensionalist Heresy.” In M. Garcia-Carpintero and J. Maciá, eds., *The Two-Dimensional Framework: Foundations and Applications* (Oxford: Oxford University Press, forthcoming).
- Stalnaker Robert. “Assertion.” *Syntax and Semantics* 9 (1978):315–32; reprinted in his *Context and Content* (New York: Oxford University Press, 1999).
- . *Inquiry*. Cambridge, MA: MIT Press, 1984.
- . “Possible Worlds.” *Noûs* 10 (1976):65–75; reprinted in Michael Loux, ed., *The Actual and the Possible* (Ithaca, NY: Cornell University Press, 1979).