

4 Proper names: Direct Reference and the Causal–Historical Theory

Overview

In a further argument against description theories of proper names, Kripke appealed to the notion of a “possible world” or universe alternative to our own. A definite description of Russell’s sort changes its referent from world to world; although “the world’s fastest woman in 1998” actually refers to Marion Jones, it designates different individuals in other worlds, since Jones might have been slower (or not even have existed) and other women might have been better runners. But typically, a proper name such as “Marion Jones” refers to the very same individual in every world in which that individual exists.

Some theorists claim that names are *directly* referential, in that a name contributes nothing but its bearer or referent to the meaning of a sentence in which it occurs. In light of Kripke’s arguments against description theories, this view is highly plausible. But the four puzzles return to haunt it. So we are left with something of a paradox.

A separate question is, in virtue of what does a proper name designate its bearer? Kripke offered a causal–historical picture of referring, according to which a given use of “Marion Jones” refers to Marion Jones in virtue of a causal chain that grounds that utterance event in the ceremony in which Jones was first given the name. But, in light of some examples that clearly do not fit that model, considerable refinement is needed to work up that picture into an adequate theory of referring.

Kripke, and Hilary Putnam, extended the causal–historical view to cover natural-kind terms, like “water,” “gold,” and “tiger,” as well as proper names. If we assume the basic correctness of that move, it has an unexpected consequence: Putnam’s famous “Twin Earth” examples seem to show that the meaning of such a term is not determined solely by what is in the heads of speakers and hearers; the state of the external world makes a contribution as well. Thus, two speakers could be molecule-for-molecule duplicates and yet mean different things by their words.

Possible worlds

I shall now set up the apparatus needed to state Kripke's fundamental criticism of description theories of proper names. I begin with the notion of a "possible world." (It goes back at least to Leibniz, though it was incorporated into philosophical logic only in the twentieth century.) Consider the world we live in—not just the planet Earth, but the whole universe. Our talk about things in our universe is talk about what actually exists, what things there really are: Gordon Brown the British Prime Minister, my left elbow, Bolivia, the sandwich on your plate, the Andromeda Galaxy, and so on, but not Hamlet, the Easter Bunny, the Big Rock Candy Mountain, or the free lunch. And what is true in this universe is of course actually true. But there are things that are in fact false, yet might have been true. Things might have gone otherwise; the world could have been different from the way it is. Someone other than Brown might have succeeded Tony Blair as Prime Minister, I might have married a different person (which would have been a mistake), and I know I could have finished writing this book sooner if I had had a private secretary and a retinue of servants including a personal chef and a couple of hired killers.

Thus there are a number of ways the world might have been. To put it slightly more fancifully, there are *alternative worlds*. Different worlds, worlds that could have been ours, but that are only possible and not actual. Think of an array of possible universes, corresponding to the infinitely many ways in which things, very broadly speaking, might have gone. All these possible worlds represent nonactual global possibilities.

Now (obviously) a sentence's truth—even when we hold the sentence's meaning fixed—depends on which world we are considering. "Brown is Prime Minister" is true in the actual world but, since Brown need not have been Prime Minister, there are countless worlds in which "Brown is Prime Minister" is false: in those worlds, Brown did not succeed Tony Blair, or never went into politics, or never even existed. And in some other worlds, someone else is Prime Minister—David Cameron, P. F. Strawson, me, Madonna, or Daffy Duck. In still others, there is no such office as Prime Minister, or not even a Britain; and so on and so forth. So a given sentence or proposition varies its truth-value from world to world.

(For now, let us take all this talk of "alternate worlds" intuitively, as a metaphor or picture, a heuristic for seeing what Kripke is getting at. Considered as serious metaphysics, it raises many controversial issues,¹ but we may hope that those issues will not much affect Kripke's use of the possible-worlds picture for his purposes in the philosophy of language.)

Just as sentences change their truth-values from world to world, a given singular term may vary its referent from world to world: in our actual world in (late) 2007, "The present British Prime Minister" designates Gordon Brown. But, as before, Brown might not have succeeded, or even gone into politics in the first place, or even existed at all. So in some other worlds the same

description, meaning what it does here in our world, designates someone else (Cameron, Strawson, . . .), or no one at all—since in some other possible worlds the Conservatives won the election, and in some there was no election, and so forth. This is why the description’s referent changes from world to world.

Let us call such a singular term, one that designates different things in different worlds, a *flaccid* designator. It contrasts specifically with what Kripke calls a *rigid* designator: a term that is not flaccid, that does not change its referent from world to world, but denotes the very same item in every world (at least in every world in which that item exists.²)

Rigidity and proper names

Now we are able to state Kripke’s further objection to description theories of proper names (1972/1980: 74ff.): a definite description of the sort Russell had in mind is flaccid, as has just been illustrated. Yet proper names, Kripke says, do not (usually) vary their reference across worlds or hypothetical situations in that way. If we imagine a world in which Aristotle does such-and-such, it is one in which *Aristotle* does that thing and has some different properties from those he has here in the real world. Our name “Aristotle” denotes *him* there, not someone else. Names are (normally) in that sense rigid designators, keeping the same referent from world to world, whereas Russellian descriptions are flaccid. Thus, names are not equivalent to Russellian descriptions. (Of course, if a description is used referentially in Donnellan’s sense, it may go rigid.)

The foregoing parenthesized qualifications (“usually,” “normally”) are important. Kripke does not hold any strict universal thesis about proper names. He is generalizing about normal uses of ordinary proper names and saying only that, for the most part, such names are used rigidly. So he is not to be refuted by coming up with unusual flaccid names, which certainly exist: occasionally, a description is offered as conventionally fixing the meaning and not just identifying the referent of an apparent proper name. “Jack the Ripper” is an example. And in popular writings about Scotland Yard or British detective culture of the 1950s, for example, the name “Chummy” was used as a mere synonym for “the culprit”; it meant, attributively or flaccidly, just “whoever committed the crime.” For that matter, probably any proper name has occasional flaccid uses. Frege (1892/1952a) offers a famous example: “Trieste is no Vienna,” where “Vienna” functions not as the name of a city, but as abbreviating a loose cluster of exciting cultural properties that Vienna has. In the same tone, on an occasion well remembered by American voters, 1988 Vice-Presidential candidate Lloyd Bentsen told his rival Dan Quayle, “Senator, you’re no Jack Kennedy.” But those are hardly standard uses of the names “Vienna” and “Jack Kennedy.”³

Kripke offers a further little intuitive test for telling whether a term is rigid: try the term in the sentence frame, “N might not have been N.” If we

plug in, for N, a description like “the President of the United States in 1970,” we obtain “The President of the United States in 1970 might not have been the President of the United States in 1970”; and the latter sentence is clearly true, at least on its most natural reading: the person who was President in 1970 might not have been President then (or at any other time). The truth of that sentence shows the description to refer to different people in different worlds, hence to be flaccid.

But if we put in the proper name “Nixon,” we get “Nixon might not have been Nixon,” at best a very strange sentence. It might mean that Nixon might not have existed at all, which is perhaps the most obvious way in which Nixon could have failed to be Nixon. But given that Nixon existed, how could he have failed to be Nixon? He could have failed to be *named* “Nixon,” but that is not to have failed to *be Nixon himself* (because, of course, Nixon need not have been named “Nixon”). He could have failed to have the properties stereotypically associated with Nixon, hence failed to “be Nixon” in the sense that Trieste fails to “be Vienna,” but as we saw in the previous chapter such flaccid uses of names are unusual.

Kripke argues that when one uses the name “Nixon” to refer to a person in this world and then starts describing hypothetical scenarios or alternative possible worlds, continuing to use the name, one is talking about the same person. So if you ask, “Might Nixon have joined the Black Panthers rather than becoming President?,” the answer may be yes or may be no, but the scenario you are considering is one in which Nixon, that very person, is a Black Panther—not one in which whoever or whatever was the U.S. President was a Panther. You are not imagining a world in which a Black Panther is President of the US.

But what of Russell’s spot-check argument? In response to “Whom do you mean by ‘Lili Boulanger’/‘Wilfrid Sellars?’” you promptly cough up a description or cluster of descriptions. Likewise Searle’s appeal to teaching and learning: They also proceed by equating the name in question with a description or cluster. These facts seem undeniable and insuperable.

In response, Kripke introduced an important distinction. Russell and Searle had both assumed that, if a name has a description or cluster associated with it in the ways they have respectively pointed out, then the name must share the meaning of that descriptive material (from now on I shall say just “description” for short). But this assumption is unwarranted, because there is a weaker relation that the description might bear to the name and still explain the spot check and pedagogical data: even though the description does not give the linguistic meaning of the name, it is what is used to determine the name’s reference on an occasion. Although the name “Lili Boulanger” is not synonymous with “the first woman ever to win the Prix de Rome,” the latter description can be used to indicate the person one is referring to when one uses “Lili Boulanger.” And it can be used as part of an explanation to a pupil, to identify the individual to which the name is attached.

Thus, even if a name in someone’s mouth at a time has a firm psychological association with a particular description in that person’s mind, it does not follow that the name is equivalent to the description in meaning. For all that has been shown, when the person obligingly coughs up the description in response to a spot check, the person is merely identifying the name’s referent. Similarly, if I tell a small child who “Gordon Brown” is, identifying that name’s referent by saying “Gordon Brown is the British Prime Minister,” it does not follow that the name “Gordon Brown” simply means “the British Prime Minister.” (Of course, this is not an argument against the Name Claim itself; it only undermines Russell’s use of the spot-check test as an argument for the Name Claim.)

Direct Reference

Russell used the four puzzles and (implicitly) his spot-check argument to attack the view that ordinary proper names are Millian names, in favor of the Description Theory. In turn, Kripke attacked the Description Theory in favor of the claim that ordinary proper names are rigid designators. But the latter claim does not quite amount to Millianism, for not all rigid designators are Millian names.

A Millian name, remember, is one that makes no propositional contribution but its bearer or referent. Its sole function is to introduce that individual into discourse; it lends nothing else to the meaning of a sentence in which it occurs. If we say “Jason is fat,” and “Jason” is a standard proper name, then the meaning of that sentence consists simply of the person Jason himself concatenated with the property of being fat.

Being Millian certainly implies being rigid. But the reverse does not hold. Although Kripke cites Mill and argues that names are rigid, rigidity does not imply being Millian. For definite descriptions can be rigid. Suppose we fall in with the prevalent view that *arithmetical* truths are all necessary truths. Then there are arithmetical descriptions, such as “the positive square root of nine,” that are rigid, because they designate the same number in every possible world, but are certainly not Millian because in order to secure their reference they exploit their conceptual content. Indeed, they seem to Russellize: “The positive square root of nine” seems to mean whatever positive number yields nine when multiplied by itself. So that description is not Millian even though it is rigid, because it does not simply introduce its bearer (the number three) into the discourse; it also characterizes three as being something that when multiplied by itself yields nine. Thus, in defending the rigidity of names, Kripke did not thereby establish the stronger claim. (Nor did he intend to; he does not believe that names are Millian.)⁴

However, other philosophers have championed the Millian conception, which has come to be called the Direct Reference theory of names. The first of these in our century was Ruth Marcus (1960, 1961), cited by Kripke as having directly inspired his work. Subsequent Direct Reference (DR) theories of

names have been built on Marcus' and Kripke's work (for example, Kaplan 1975; Salmon 1986).

The latter theorists have extended DR to cover some other singular terms, notably personal and demonstrative pronouns such as "I," "you," "she," "this," and "that," as well as names. (An obvious problem about extending DR to pronouns is that any normal speaker of English knows what they mean, whether or not the speaker knows whom they are being used to designate on a given occasion; if you find "I am ill and will not hold class today" written on the blackboard in an empty classroom, you understand the sentence even if you do not know who wrote it or on what day. This problem will be addressed in chapter 11.)

Of course, DR must confront the four puzzles. And, obviously, the DR theorist cannot subscribe to Russell's solution or anything very like it, for, according to DR, names do nothing semantically but stand for their bearers.

Let us reconsider the Substitutivity puzzle first. Recall our sentence:

- (1) Albert believes that Samuel Langhorne Clemens was less than 5 feet tall.

(1) goes false when "Mark Twain" is substituted for "Samuel Langhorne Clemens." How can DR explain or even tolerate that fact?

DR theorists employ a two-pronged strategy. There is a positive thesis and there is a negative thesis (though these are not often explicitly distinguished). The positive DR thesis is that the names in question really do substitute without altering the containing sentence's truth-value. On this view,

- (2) Albert believes that Mark Twain was less than 5 feet tall.

is true, not false. At the very least, belief sentences have transparent readings or understandings, on which readings the names that fall within the scope of "believes" really do just refer to what they refer to.

We naturally think otherwise; (2) does not seem true to us. That is because when we see a belief sentence, we usually take its complement clause to reproduce the ways in which its subject would speak or think. If I assert (2), I thereby somehow imply that Albert would accept the sentence "Mark Twain was less than 5 feet tall" or something fairly close to it. If I say, "Albert doesn't believe that Mark Twain was less than 5 feet tall," I am suggesting that if faced with the sentence "Mark Twain was less than 5 feet tall," Albert would say either "No" or "I wouldn't know."

But the DR theorists point out that such suggestions are not always true, perhaps not ever true. Consider:

- (3) Columbus believed that Castro's island was China.

(Attributed to the late Roderick Chisholm.) We all know what one would mean in asserting (3); the speaker would mean that when Columbus sighted

Cuba he thought that he was nearing the East Indies and was directly approaching China. Of course, being 450 years early, Columbus did not know anything about Fidel Castro; yet we can assert (3) with no presumption that its complement clause represents things in the way that Columbus himself represented them. *The speaker* makes this reference to Cuba without at all assuming that Columbus would have referred to Cuba in that way or in any parallel or analogous way.

Or suppose you and I are among the few people who know that our acquaintance Jacques is in fact the notorious jewel thief who has been terrorizing Paris' wealthy set, called "Le Chat" in the popular press and by the gendarmes. We read in the newspaper after a particularly daring but flawed robbery that police believe "Le Chat dropped the fistful of anchovies as he or she ran." We say to each other, "The police think Jacques dropped the anchovies as he ran."

So it seems undeniable that there are transparent positions inside belief sentences, in which the referring expression does just refer to its bearer, without any further suggestion about the way in which the subject of the belief sentence would have represented the bearer. Singular terms can be and are often understood transparently. We might even say:

(4) Some people doubt that Tully is Tully.

meaning that some people have doubted of the man Cicero that he was also Tully. That would perhaps be a minority interpretation of (4), but we *can* at least hear (4) as asserting that the people doubt of Cicero that he was Tully.⁵

Virtually all the DR literature has been devoted to establishing the positive thesis, that names do have Millian readings even in belief contexts. But the positive thesis is far from all that the DR theorist needs. For, although we may be persuaded that every belief sentence does have a transparent reading, most of us also remain convinced that every belief sentence also has an opaque reading, that on which some substitutions turn truths into falsehoods: *in one sense* Columbus believed that Castro's island was China, but in another, he believed no such thing, for the obvious reason that he had never heard (and would never hear) of Castro. Similarly, in one sense the police believe that Jacques dropped the anchovies, but in another they do not, and likewise for people doubting "that Tully is Tully." Yet it seems DR cannot allow so much as a sense in which belief contexts are opaque. That is DR's negative thesis: that names do not have non-Millian readings, even in belief contexts.

The problem gets worse: it is hard to deny that the opaque readings are more readily heard than the transparent readings. Indeed, that is implicitly conceded by the DR theorists, in that they know they have had to work to make us hear the transparent readings. The DR theorists must try to explain the fact away as a particularly dramatic illusion. That is, they must hold that, in fact, sentences like (1)–(4) cannot literally mean what we can and usually would take them to mean; there is some extraneous reason why we are seduced into hearing such sentences opaquely. A few such putative explanations have

been sketched, using materials we shall encounter in chapter 13 (Salmon 1986; Soames 1987, 2002; Wettstein 1991; and see Marcus 1981). But here, in my opinion, the DR theorists have come up short; at least, none of the sketches produced to date has struck me as very plausible, though perhaps Soames (2002) is the most promising.

As is implied by example (4), Frege's Puzzle is even worse for the Millian. According to DR, a sentence like "Samuel Langhorne Clemens is Mark Twain" can mean only that the common referent, however designated, is himself. Yet such a sentence is virtually never understood as meaning that. And anyone might doubt that Clemens is Twain, seemingly without doubting anyone's self-identity. Here again, the DR theorist bears a massive burden, of explaining away our intuitive judgements as illusory.

The Problems of Apparent Reference to Nonexistents and Negative Existentials are if anything worse yet. If a name's meaning is simply to refer to its bearer, then what about all those perfectly meaningful names that have no bearers? (But for doughty recent DR attempts to meet those two problems, see Salmon 1998; Soames 2002; Braun 2005.)

We have come to a deep dilemma, nearly a paradox. On the one hand, in chapter 3 we saw compelling Kripkean reasons why names cannot be thought to abbreviate flaccid descriptions, or otherwise to have substantive senses or connotations. Intuitively, names are Millian. Yet, because the original puzzles are still bristling as insistently as ever, it also seems that DR is pretty well refuted. This is a dilemma, or rather trilemma, because it further appears that we are stuck with one of those three possibilities: either the names are Millian, or they abbreviate descriptions outright, or in some looser way such as Searle's they have some substantive "sense" or content. But none of these views is acceptable.

A few theorists have claimed to find ways between the three horns. As noted in chapter 3, Plantinga (1978) and Ackerman (1979) have appealed to rigidified descriptions, such as "the *actual* winner of the 1968 election" as opposed to just "the winner of the 1968 election"; the former description is rigid, because "the actual winner" means the winner in this (our) world, and refers to that same person at any other world regardless of whether he won there. On this view, names are rigid though they are not Millian. (See also Jackson 1998.) Michael Devitt (1989, 1996) has offered a radical revision of Frege's notion of sense. I myself (Lycan 1994) have offered a much subtler, more beautiful and more effective weakened version of DR, but it would be immodest of me to tout it here.⁶

We must now make a crucial distinction. So far in this chapter we have been talking about the *semantics of proper names*, that is, about theories of what names contribute to the meanings of sentences in which they occur. DR in particular takes for granted the idea of a name's referent or bearer. But then a separate question is. In virtue of what is a thing the referent or bearer of a particular name? Semantics leaves that question to philosophical analysis. A *philosophical theory of referring* is a hypothesis as to what relation it is exactly

that ties a name to its referent—more specifically, an answer to the question of what it takes for there to be a referential link between one’s utterance of a name and the individual that gets referred to by that utterance.

Semantical theories of names and philosophical accounts of referring vary independently of each other. The difference was blurred by Russell and by Searle,⁷ because each gave a very similar answer to both questions. Russell said that a name gets its meaning, and contributes to overall sentence meaning, by abbreviating a description; *also*, what makes a thing the bearer of the name is that the thing uniquely satisfies the description. Likewise for Searle and his clusters. But now notice that, if one is a DR theorist, that alone tells us nothing about what attaches a name to its referent. The same goes for Kripke’s weaker rigidity thesis; up till now, he has been talking semantics only, and we have seen nothing of his theory of referring. To that we now turn.

The Causal–Historical Theory

As you can verify for yourself, most of Kripke’s objections to the Name Claim and to description semantics generally will also translate into objections to the Description Theory of referring; the Description Theory will predict the wrong referent (think of the Gödel/Schmidt example in objection 5, chapter 3) or no referent at all (as when there is no particular description the speaker has in mind (objection 1) or in indefinite cases, as in objection 6).

Kripke sketches a better idea. He begins memorably (1972/1980: 91): “Someone, let’s say, a baby, is born” (I think we should grant Kripke’s assumption that the neonate is a baby. There is such a thing as being too picky.) He continues:

[The baby’s] parents call him by a certain name. They talk about him to their friends. Other people meet him. Through various sorts of talk the name is spread from link to link as if by a chain. A speaker who is on the far end of this chain, who has heard about, say Richard Feynman, in the market place or elsewhere, may be referring to Richard Feynman even though he can’t remember from whom he first heard of Feynman or from whom he ever heard of Feynman. He knows that Feynman was a famous physicist. A certain passage of communication reaching ultimately to the man himself does reach the speaker. He then is referring to Feynman even though he can’t identify him uniquely.

The idea, then, is that my utterance of “Feynman” is the most recent link in a causal–historical chain of reference borrowings, whose first link is the event of the infant Feynman’s being given that name. I got the name from somebody who got it from somebody else who got it from somebody else who got it from somebody else . . . , all the way back to the naming ceremony. I do not have to be in any particular cognitive state of Russell’s or Searle’s

sort. Nor need I have any interesting true belief about Feynman, or as to how I acquired the name. All that is required is that a chain of communication in fact has been established by virtue of my membership in a speech community that has passed the name on from person to person, which chain goes back to Feynman himself.

Of course, when a new user first learns a name from a predecessor in the historical chain, it can only be by the newbie's and the predecessor's sharing a psychologically salient backing of identifying descriptions. But, as before, there is no reason to assume that that particular backing of descriptions fixes the name's sense. It is needed only to fix reference. So long as the newbie has a correct identificatory fix on the predecessor's referent, the newbie can then freely use the name to refer to that person.

Taken at face value, this causal–historical view makes the right predictions about examples such as Donnellan's Tom. In each example, referring succeeds because the speaker is causally connected to the referent in an appropriate historical way.

Kripke (1972/1980: 66–7) offers the further case of the biblical character Jonah. It is similar to the “Nixon” example (objection 3 in chapter 3). Kripke points out that we should distinguish between stories that are complete legends and stories that are, rather, substantially false accounts of real people. Suppose historical scholars discover that in fact no prophet was ever swallowed by a big fish, or did anything else attributed by the Bible to Jonah. The question remains of whether the Jonah character was simply made up in the first place, or whether the story is grounded ultimately in a real person. Actually there are subcases: someone could have made up and spread a host of false stories about Jonah immediately after his death; or because Jonah was an exciting individual, all sorts of rumors and stories began to circulate about him, and the rumors got out of hand; or there might have been a very gradual loss of correct information and accretion of false attributions over the centuries. But, in any of these cases, it seems that today the Bible is saying false things about the real person, Jonah.⁸

It might be thought that *ambiguous* names—names borne by more than one person—pose a problem for the causal–historical view. (“John Brown” is ambiguous as between the former Scots ghillie who befriended Queen Victoria after Albert's death, the monomaniacal failed farmer who invaded Harper's Ferry in 1859, and doubtless thousands of other males of the English-speaking world. Until 1994, even the highly distinctive name “William Lycan” applied to more than one person. I suppose the vast majority of names are ambiguous; a name is unambiguous only by historical accident.) This is no problem at all for description theories because, according to description theories, ambiguous names simply abbreviate different descriptions. (If anything, description theories make proper names *too richly* ambiguous.) But what if you endorse DR, and you deny that names have senses or descriptive connotations in any sense at all?

I asked that last question only to see if you had been paying attention earlier. It flagrantly ignores the important distinction between the semantics of names and the theory of referring. The Causal–Historical Theory of referring has a straightforward answer to the question of ambiguous names: if a name is ambiguous, that is because more than one person has been given it. What disambiguates a particular use of such a name on a given occasion is—what else?—that use’s causal–historical grounding, specifically the particular bearer whose naming ceremony initiated its etiology.

Kripke emphasizes that he has only sketched a picture; he does not have a worked-out theory. The trick will be to see how one can take that picture and make it into a real theory that resists serious objections. The only way to make a picture into a theory is to take it overliterally, to treat it as if it were a theory and see how it needs to be refined. Kripke does just that, though he leaves the refinement to others.

Problems for the Causal–Historical Theory

The causal–historical view’s key notion is that of the passing on of reference from one person to another. But not just any such transfer will do. First, we must rule out the “naming after” phenomenon. My boyhood friend John Lewis acquired a sheepdog, and named it “Napoleon” after the emperor; he had the historical Napoleon explicitly in mind and wanted to name his dog after that famous person. “Naming after” is a link in a causal–historical chain: it is only because the emperor was named “Napoleon” that John Lewis named his dog that. But it is the wrong kind of link. To rule it out, Kripke requires that “[w]hen the name is ‘passed from link to link’, the receiver of the name must . . . intend when he learns it to use it with the same reference as the man from whom he heard it” (1972/1980: 96). This requirement was clearly not met by John Lewis, who was deliberately changing the referent from the emperor to the dog and meant his friends to be well aware of that.

Second, Kripke adduces the example of “Santa Claus.” There may be a causal chain tracing our use of that name back to a certain historical saint, probably a real person who lived in eastern Europe centuries ago, but no one would say that when children use it they unwittingly refer to that saint; clearly they refer to the fictional Christmas character. But then, how does “Santa Claus” differ from “Jonah”? Why should we not say that there was a real Santa Claus, but that all the mythology about him is garishly false? Instead, of course, we say that there is no Santa Claus (apologies to anyone who did not know that). We use the name “Santa Claus” as though it abbreviates a description. A similar example would be that of “Dracula.” It is well known that the contemporary use of that name goes back to a real Transylvanian nobleman called “Vlad” (commonly, “Vlad the Impaler,” in virtue of his customary treatment of people who had annoyed him). But of course when we now say “Dracula” we mean the fictional vampire created by Bram Stoker and portrayed by Bela Lugosi in the famous movie.

Having merely raised the problem, Kripke does not try to patch his account in response, but moves on. Probably the most obvious feature to note is that “Santa Claus” and “Dracula” as we use those names are associated with very powerful stereotypes, indeed cultural icons in the United States. Their social roles are so prominent that they really have ossified into fictional descriptions, in a way that “Jonah” has not even among religious people. In a way, Jonah’s iconic properties are side by side with his historical properties in the Old Testament, but as we might say, “Santa Claus” and “Dracula” are pure icons. And for the average American, the myth utterly dwarfs the historical source.

As Kripke says, much work is needed. Devitt (1981a) offers a fairly well developed view that does qualify as a theory rather than only a picture. However, here are a few objections that would apply to any version of the Causal–Historical Theory as described above.

OBJECTION 1

We have been offered the notion of a causal–historical chain leading back in time from our present uses of the name to a ceremony in which an actual individual is named. But how, then, can the Causal–Historical Theorist accommodate empty names, names that have no actual bearers?

Perhaps the best bet here is to exploit the fact that even empty names are introduced to the linguistic community at particular points in time, either through deliberate fiction or through error of one kind or another. From such an introduction, as Devitt (1981a) and Donnellan (1974) point out, causal–historical chains begin spreading into the future just as if the name had been bestowed on an actual individual. So reference or “reference” to nonexistent is by causal–historical chain, but the chain’s first link is the naming event itself rather than any putative doings of the nonexistent bearer.⁹

OBJECTION 2

Evans (1973) points out that names can change their reference unbeknownst, through mishap or error, but the Causal–Historical Theory as presented so far cannot allow for that. According to Evans,¹⁰ the name “Madagascar” originally named, not the great African island, but a portion of the mainland; the change was ultimately due to a misunderstanding of Marco Polo’s. Or:

Two babies are born, and their mothers bestow names upon them. A nurse inadvertently switches them and the error is never discovered. It will henceforth undeniably be the case that the man universally known as “Jack” is so called because a woman dubbed some other baby with the name.

(Evans 1973: 196)

We do not want to be forced to say that our use of “Madagascar” still designates part of the mainland, or that “Jack” continues to refer to the other former baby rather than to the man everyone calls “Jack.”

In reply, Devitt (1981a: 150) suggests a move to *multiple grounding*. A naming ceremony, he says, is only one kind of occasion that can ground an appropriate historical chain; other perceptual encounters can serve also. Instead of there being just the single linear causal chain that goes back from one’s utterance to the original naming ceremony, the structure is mangrove-like: the utterance proceeds also out of further historical chains that are grounded in later stages of the bearer itself. Once our use of “Madagascar” has a large preponderance of its groundings in the island rather than the mainland region, it thereby comes to designate the island; once our use of “Jack” is heavily grounded in many people’s perceptual encounters with the man called that, those groundings will overmaster the chain that began with the naming ceremony. This is vague, of course, perhaps objectionably so.

OBJECTION 3

We can misidentify the object of a naming ceremony. Suppose I am seeking a new pet from the Animal Shelter. I have visited the Shelter several times and noticed an appealing grey tabby; I decide to adopt her. On my next visit I prepare to name her. The attendant brings out a tabby of similar appearance and I believe her to be the same one I intend to adopt. I say, “Here we are again, then, puddy-tat; your name is now ‘Liz’, after the composer Elizabeth Poston, and I’ll see you again after you’ve had all your shots” (tactfully I do not mention the mandatory neutering). The attendant takes the cat away again. But unbeknownst to me it was the wrong cat, not my intended pet. The attendant notices the mistake, without telling me, recovers the right cat, and gives her her shots (and the rest). I pick her up and take her home, naturally calling her “Liz” ever thereafter.

The problem is of course that my cat was not given that name in any ceremony. The imposter was given it, even if I had no right to name her. Yet surely my own cat is the bearer of “Liz,” not just after subsequent multiple groundings have been established, but even just after the naming ceremony I did perform. (It would be different if I had taken the imposter home and continued to call *her* “Liz.”) The multiple-grounding strategy does not seem to help here. Rather, what matters is which cat I *had in mind* and believed I was naming in the ceremony. (Devitt (1981a: section 5.1) speaks of “abilities to designate,” construing these as mental states of a certain sophisticated type.) If so, then repair of the Causal–Historical Theory on this point will require a significant foray into the philosophy of mind.

OBJECTION 4

People can be *categorically* mistaken in their beliefs about referents. Evans cites E. K. Chambers' *Arthur of Britain*¹¹ as asserting that King Arthur had a son Anir "whom legend has perhaps confused with his burial place." A speaker in the grip of the latter confusion might say "Anir must be a green and lovely spot"; the Causal–Historical Theory would interpret that sentence as saying that a human being (Arthur's son) was a green and lovely spot. Less dramatically, one might mistake a person for an institution or vice versa. (A former colleague of mine used to use the name of Emerson Hall—the building that houses the Harvard philosophy department—as a way of referring to the department, as in "Emerson Hall isn't going to like this." A casual hearer might easily have gotten the idea that "Emerson Hall" names a person.) Or one might mistake a shadow for a live human being and give it a name. In none of these cases is it plausible to say that subsequent uses of the name in question really refer to the categorially erroneous item.

Devitt and Sterelny (1987) call this the "qua-problem." They concede that the celebrant at a naming ceremony, or other person responsible for any of a name's groundings, must not be categorically mistaken and must indeed intend to refer to something of the appropriate category. This is a mild concession to Descriptivism.

There are more objections (some of them further ones of Evans'). The majority position seems to be that Kripke initially overreacted to the Descriptivist picture. He was right to insist that causal–historical chains of some kind are required for referring and that descriptions do not do nearly as much work as Russell or even Searle thought they did; but (as critics, including Kripke himself, maintain) there still are some descriptive conditions as well. The trick is to move back in the direction of Descriptivism without going so far as even Searle's weak Descriptivist doctrine. But that does not leave much room in which to maneuver. One promising line has come to be called "causal descriptivism" (Kroon 1987): The idea is to accept the causal–historical theorist's causal–historical proposal as getting the cases right, but make it into a descriptive condition. Kroon defends this primarily by extending the "naming after" objection.

Natural-kind terms and "Twin Earth"

Kripke (1972/1980) and Hilary Putnam (1975a) went on to extend both the semantic theory of rigid designation and the Causal–Historical Theory of referring from singular terms to some predicates or general terms, chiefly *naturalkind terms*, common nouns of the sort that refer to natural substances or organisms, like "gold," "water," "molybdenum," "tiger," and "aardvark." Such expressions are not singular terms, since they do not purport to apply to just one thing. But Kripke and Putnam argued that they are more like names than they are like adjectives. Semantically they are rigid; each refers

to the same natural kind in every world in which that kind has membership. And some version of the Causal–Historical Theory characterizes their referring use.

This view sharply opposed a long-held Descriptivist theory of natural kind terms, which associated each such term with a descriptive stereotype. For example, “water” would have been analyzed as meaning something like “a clear, odorless, tasteless potable liquid that falls from the sky as rain and fills lakes and streams,” and “tiger” as something like “a ferocious, carnivorous jungle feline, tawny with distinctive black stripes.” Kripke and Putnam urged modal arguments against such analyses, similar to objection 3 from the previous chapter and to the rigidity argument that began this one. For example, there could have been water even if there had never been rain, lakes or streams, and under different circumstances water might have had an odor or a taste. Tigers might have been born tame, and we might even find out that no tiger has ever in fact had stripes (a worldwide Wonderland-style conspiracy might have had all the stripes painted on).

What does make something a tiger, then, or a sample of water, if not the commonsense stereotype? Kripke and Putnam adverted to the scientific natures of natural kinds. What makes water water is its chemical composition, that it is H_2O ; what makes tigers tigers is their distinctive genetic code. In every possible world, water is H_2O , but in some worlds H_2O has an odor or a taste.

It may be objected that the chemical composition of water and the genetic characteristics of tigers were highly substantive empirical discoveries; so surely it was possible that water not be H_2O , and so there are worlds in which water is not H_2O . But Kripke and Putnam rejoined that the alleged “possibility” here is only a matter of scientific ignorance, not a genuine metaphysical possibility; once one does discover the scientific essence of a natural kind, one has discovered the true metaphysical nature of that kind, and the kind has that nature in every possible world in which it is manifested. What change from world to world are the elements of the commonsense stereotype.

If this view is correct,¹² it has a somewhat startling implication about the relation between linguistic meaning and the mind: as Putnam puts it, that meaning “ain’t in the head.” Putnam imagines that somewhere in another galaxy there is a planet, called “Twin Earth,” which is a nearly exact duplicate of our Earth, running along exactly in parallel with our own history. It contains a Twin Putnam, a Twin Brooklyn Bridge, a Twin Lycan and a Twin you, all molecular duplicates of their counterparts here. If one were able to watch both planets simultaneously, it would be like watching the same television program on two different screens. (But it is important to note that Twin Earth is not a different possible world; it is only another planet, within the same world as Earth. Though exactly like you to look at and embedded in an almost exactly similar planetary context, of course Twin you is not you, but a numerically different person.)

I said that Twin Earth is a *nearly* exact duplicate of Earth. There is one difference: what looks and behaves like water on Twin Earth is not water—that is, H_2O —but a different substance that Putnam calls XYZ. XYZ is odorless and tasteless and has the other superficial properties that water does, but it is only “fool’s water” (as in “fool’s gold”). Of course, the Twin-English-speaking Twin Earthlings call the XYZ “water,” since they are otherwise just like us,¹³ but that is an equivocation; “water” in Twin English means XYZ, not water, just as (I am told) the kind term “chicory” in British English means a different plant from the one meant by the same word in American English.

Now, consider a pair of transworld twins, say Gordon Brown and Twin Gordon. After a natural disaster, Brown emphasizes the urgency of getting food and water to the victims. Naturally, at the same time, Twin Gordon emphasizes the urgency of getting food and “water” to the victims. But the word-for-word-identical sentences they utter mean different things. Brown’s sentence means that the victims must be provided with food and H_2O , while Twin Gordon’s means that the victims must be provided with food and XYZ.

Yet Brown and Twin Gordon are physical duplicates. Given Putnam’s background assumptions, this shows that the meanings of Brown’s and Twin Gordon’s utterances are not determined by the total states of their brains, or even by the total states of their bodies. For their brain states and bodily states are identical, yet their utterance meanings differ.

That is perhaps no big surprise. After all, language is public property; any given language is used by a community, for communication between different people, not often for the mere articulation of someone’s private thoughts. But in fact (again given the background assumptions), Putnam’s example shows more: that the linguistic meanings of sentences are not determined even by the *totality* of speakers’ brain states and bodily states, indeed even by the entire community’s pattern of usage. For English speakers and Twin English speakers are all exactly alike in their physical composition and in the public deployment of words that sound just the same; yet the sentences of their otherwise identical languages mean different things.¹⁴ We shall return to this point in chapter 6.

Now it is time to branch out and take up the whole matter of meaning and theories of meaning.

Summary

- Kripke argued that proper names normally function as rigid designators, that a name denotes the same individual in every possible world in which that individual exists.
- Taking a more ambitious line, the DR theorists defend the Millian view that a name’s sole contribution to the meaning of a sentence in which it occurs is to introduce its bearer into the discourse.

- But our four logical puzzles about reference still arise just as insistently as before, and seem to make DR untenable. We are left with something of a paradox.
- Turning to the theory of referring, Kripke offered his causal–historical picture as a replacement for Description Theories. Michael Devitt and others have refined and ramified the causal–historical view in response to initial objections.
- Kripke and Putnam extended the Causal–Historical Theory to cover natural-kind terms.
- If the Causal–Historical Theory is correct, then Putnam’s “Twin Earth” examples seem to show that the meanings of a speech community’s words are not entirely determined by the contents of speakers’ and hearers’ heads; the external world makes a contribution also.

Questions

- 1 Some philosophers are uneasy about Kripke’s notion of a “rigid designator” and his ancillary distinction between “fixing sense.” If you too are uneasy about “rigidity,” can you articulate the problem?
- 2 Do fictional names pose a special problem for Kripke’s rigidity thesis? How might he treat fictional names?
- 3 Now that Kripke has rejected the Name Claim, how might he address one or more of the four puzzles?
- 4 Can you help DR address one or more of the four puzzles (a harder task)?
- 5 Can you respond more fully on behalf of the Causal–Historical Theory to objections 1–4?
- 6 Offer your own criticisms of the causal–historical picture.
- 7 Assess the Kripke–Putnam view that natural-kind terms rigidly designate scientifically characterized kinds.
- 8 Are you persuaded by Putnam’s “Twin Earth” examples that meanings “ain’t in the head”?

Further reading

- Further representative papers on Direct Reference may be found in Almog, Perry and Wettstein (1989); a survey and critique is offered in Devitt (1989). See also Recanati (1993).
- Kvart (1993) also elaborates a version of the Causal–Historical Theory of referring.
- Evans (1973) offers further objections to Kripke’s picture, and an interesting revision of it. Evans (1982) makes concessions to Kripke but insists that the idea of a “name-using (social) practice” must be introduced as a further element. McKinsey (1976, 1978) has moved

further back toward the *ancien régime*. Further objections are made by Erwin, Kleiman and Zemach (1976) and Linsky (1977).

- Salmon (1981) surveys semantic views of kind terms. Schwartz (1977) contains relevant papers. Criticisms of the Kripke–Putnam line are offered by Fine (1975), Dupré (1981), Unger (1983), and others. Boër (1985) responds to some of those criticisms.
- The impact of “Twin Earth” examples on the theory of meaning generally is explored in Harman (1982) and Lycan (1984: ch. 10).