

Fruits of the Causal Theory of Reference

Scott Soames

The causal theory of reference arose from Saul Kripke's attack in *Naming and Necessity* on descriptivist analyses of proper names. His target was the view that, like most meaningful expressions, proper names express concepts that determine their extensions. In the case of names, these were thought to be individual concepts the unique instantiators of which were their referents. Since singular definite descriptions also express such concepts, it seemed obvious that names must be synonymous with descriptions associated with them by users. Kripke's attack on this view was an earthquake that shook the foundations of natural-language semantics -- despite the fact that proper names plays a very small role in the overall enterprise.

It did so because it challenged a fundamental tenet of the then dominant conception of semantic theorizing, summarized as follows:

The Meaning of a Non-Indexical Expression E is, (a) what is understood by speakers who use E competently and so know what E means, (b) what E contributes to the compositionally determined meanings (semantic contents) of all sentential clauses in which E occurs, (c) what determines the extension of E at world-states, thereby contributing to modal truth conditions of sentences containing E, and (d) what E standardly contributes to illocutionary contents of uses of sentences containing E.

When Kripke concluded that the meanings of names are never those of descriptions, while refusing to speculate about what names do mean, he rejected (c) above, arguing that the referent of a name at any world-state is fixed by the causal-historical chain associated with it at the actual world-state. When he extended this view to natural kind terms, the repudiation of standard semantic architecture became more serious and was resisted by causal descriptivists who sought to preserve it by incorporating the causal-historical theory of reference into their analysis names and natural kind terms. The story of the resulting dispute and both the positive

insights and the further challenge to traditional semantic architecture to which it led is the story of the fruits of the causal theory of reference.

The chief strategy for defending descriptivist theories of the meaning of names and natural kind terms was to rigidify alleged reference-fixing descriptions, thereby avoiding Kripke's modal argument. The chief strategy for defending descriptivist accounts of reference was to invoke descriptions other than those Kripke had shown to be untenable. These defenses were motivated by a striking idea; since facts, not magic, determine what a term refers to, it must be possible to formulate and verify a theory giving those facts.¹ Since verification requires showing the referents of a term *t* at different world-states match what we, ordinary users, would take *t* to refer to (as used there), a description extractable from the theory we implicitly follow, typically without being able to explicitly formulate it, *must* correctly fix *t*'s referent. In short, descriptivism, properly understood, can't possibly fail.

Although influential, this argument it provoked suspicion. How could what is, in effect, an apriori argument tell us what form an empirical theory must take? Before addressing this perplexity, we must note another factor motivating resistance to Kripke. Millianism, which was the theory of meaning most congenial to his antidescriptivism, seems incomplete, problematic, or both. Surely, one is inclined to think, the referents of 'Hesperus', 'Phosphorus', and 'Earth', don't exhaust their meaning; somehow, more content must be included. Moreover, the assertions made, and beliefs expressed, using the names don't seem to survive when different coreferential names are used. Doesn't this show that the meanings of names are often not their referents, and may sometimes be descriptive? Isn't the case even stronger for natural kind terms? Surely, there is more to "knowing what they mean," than simply being able to use them

¹ Frank Jackson puts things this way on p. 82 of Jackson (1998a).

the designate the relevant kinds. Understanding the causal theory of reference, and its significance, requires addressing these concerns.

The Anti-Descriptivist Treatment of Names in *Naming and Necessity*

Kripke distinguished two versions of descriptivism -- one purporting to give the meaning and referent of a proper name for a speaker and one purporting only to do the latter.

Strong Descriptivism about Names

The meaning of a name *n* for a speaker *x* is a description *D* associated with *n* by *x*; its referent at *w* is whatever uniquely satisfies *D* at *w*.

Corollary: If *D* gives the meaning of *n*, then the propositions semantically expressed by $\lceil n \text{ is } F \rceil$, $\lceil D \text{ is } F \rceil$ are the same, as are those expressed by $\lceil \text{if } n \text{ exists, then } n \text{ is } D \rceil$, $\lceil \text{if } D \text{ exists, } D \text{ is } D \rceil$. Since the latter is both necessary and knowable apriori, so is the former. In general, substitution of *D* for *n* in a sentence *S* preserves modal and epistemic profile.²

Weak Descriptivism about Names

Descriptions associated with a name by speakers semantically *fix its referent* without giving its meaning. Once fixed at the actual world state, it remains the same at other world-states.

Corollaries: (i) Speakers believe they can uniquely describe what they use names to refer to; (ii) When *D* semantically fixes the referent of *n*, then *n* refers to *o* iff *o* uniquely satisfies *D*. (iii) When *D* fixes *n*'s referent for *A*, *A* knows on the basis of semantic knowledge alone that if $\lceil n \text{ exists (or existed)} \rceil$ expresses a truth, then $\lceil n \text{ is (or was) } D \rceil$ also expresses a truth.

If *D* fixes the referent of *n*, $\lceil n \text{ is } F \rceil$ and $\lceil D \text{ is } F \rceil$ express different propositions. $\lceil n \text{ is } F \rceil$ expresses a singular proposition analogous that expressed by a use of $\lceil \text{that is } F \rceil$, said demonstrating *o*.³

Against Strong Descriptivism

Having distinguished these versions of descriptivism, Kripke observed that proper names are rigid designators, even though the descriptions speakers typically associate with them -- which, he seemed to assume, are those they would most readily volunteer if asked "To

² 'D' and 'n' are metalinguistic variable. The square quotes are called "corner" or "Quine" quotes. E.g.,. If 'P' and 'Q' are variables over sentences, the sentence *For all sentences P and Q $\lceil P \ \& \ Q \rceil$ is a sentence* says *For all sentences P and Q, the expression that consists of P, followed by '&', followed by Q is a sentence*. Similarly, *For all sentences P, $\lceil P \text{ is true iff } P \rceil$ says* *For all sentences P, the expression that consists of the left hand quote mark, followed by P, followed by the right hand quote mark, followed by 'is true iff,' followed by P is true*.

³ Page 57 of Kripke (1980).

whom/what are you using n to refer?" -- aren't. This is important because rigid and non-rigid designators make different contributions to the modal truth conditions of sentences containing them. If I were to say "Aristotle could have been so-and-so," there is a single individual whose *being so-and-so* at any possible world-state is sufficient for the truth of my remark and whose *being so-and-so* at some possible state is necessary for its truth. Since 'Aristotle' is rigid, the needed verifier doesn't change from world-state to world-state.

Kripke puts this by saying that a use of 'Aristotle' (at the actual world-state) designates the same man at any world-state at which he exists, and never designates anything else. This isn't true of most well-known descriptions we associate with the name -- e.g., 'the greatest student of Plato', 'the father for formal logic,' or 'the teacher of Alexander the Great'. Because they aren't rigid, we can say that had the world been different, Aristotle would have existed without being the greatest student of Plato, the father of formal logic, or the teacher of Alexander. Since we can't correctly say that had the world been different Aristotle would have existed without being Aristotle, none of those descriptions, no conjunction or disjunction of them and no nonrigid description generally is synonymous with 'Aristotle'. That, in brief, is Kripke's modal argument.

Can strong descriptivism be saved by rigidifying the putative meaning-giving descriptions? One way of rigidifying uses the indexical operator *actually* of two-dimensional modal logic. If S expresses proposition p at a context C (and an assignment of values to variables) $\lceil \text{Actually } S \rceil$ expresses the proposition that p is true at the world-state @ of C. So, when S is true at @ $\lceil \text{Actually } S \rceil$ is necessary. When a nonrigid description $\lceil \text{the } x: Gx \rceil$ uniquely designates o at @, the description $\lceil \text{the } x: \text{Actually } Gx \rceil$ rigidly designates o at all world-states possible from @ at which o exists and never designates anything else.

Thus, if Kripke's modal argument were the only problem for strong descriptivism, rigidifying descriptions in this way would save it. But there are other problems. Just as substituting nonrigid descriptions for names can change the *modal profile* of a sentence S, substituting 'actually'-rigidified descriptions can change the *epistemic profile* of S. This, in turn, leads to different modal-profiles of belief reports differing only in such substitution.

This is illustrated in (1), where 'the x: Gx' nonrigidly designates the bearer of name 'n'.

- 1a. Bill believes that n is F
- b. Bill believes that the x: Gx is F
- c. Bill believes that the x: actually Gx is F

Since 'n' rigidly designates o, the belief (1a) reports Bill as having will be true at w only if o is F at w. Because the description in (1b) isn't rigid, the belief it reports doesn't have this property. Since (1a,b) report different beliefs, 'n' doesn't mean what 'the x: Gx' does. (1c) also differs from (1a). For the belief reported by our use of (1c) to be true at w, Bill must believe (at w) that *the unique individual who, at @, is G is F* (at w), which is not required by (1a). Since substitution changes the epistemic profile of $\lceil n \text{ is F} \rceil$, $\lceil \text{the x: actually Gx is F} \rceil$ doesn't give the meaning of n. Thus, names don't mean the same as 'actually'-rigidified descriptions.

Suppose, however, that a potentially meaning-giving description D is rigidified using the indexical *dthat* operator of Kaplan (1989), which, when prefixed to D produces a singular term the *semantic content* of which is the object denoted by D. In one way, this is just what the doctor ordered, since *if genuine reference-fixing descriptions of the kind sought by weak descriptivism can be found*, then rigidifying with 'dthat' will avoid the modal argument, without succumbing to this objection.⁴ Still the price of this move is steep, since, without other

⁴ On this analysis, strong descriptivism maintains that both the meanings and semantic contents of names are the same as the meanings and semantic contents of *dthat-rigidified descriptions*, though since both names and such descriptions are taken to be indexical, the meanings and semantic contents of each are distinct.

potentially problematic semantic changes, rigidifying with 'dthat' reinstates Frege's Puzzle, the avoidance of which has, historically, been the chief force motivating descriptivism.

Against Weak Descriptivism

Unlike strong descriptivism, the purportedly reference-fixing descriptions associated with names by weak descriptivism needn't also determine their meanings. Hence, it is not defeated by the fact that substituting them for names may change the modal or epistemic profile of a sentence. Nevertheless, Kripke argued, it is refuted by arguments showing that, for most names (and natural kind terms), no semantic rule determines their reference to be whatever, if anything, uniquely satisfies descriptions speakers associate with them. Rather, we often successfully use a name to refer to its conventional bearer, even though the information we associate with it is incomplete or erroneous. For example, speakers who know of Cicero only that he was a famous Roman, or of Einstein only that he was a brilliant scientist, can use the names to refer to the men, despite realizing that they can't describe them uniquely. Thus, it seems, the referent of their use of the names can't be determined by any (uniquely identifying) description they associate with them.

In other cases, one's use of *n* refers to its bearer despite being guided by a description that denotes something else. Imagine a student whose use of 'Thales' is guided by the description "the pre-Socratic who held that all is water." Suppose some pre-Socratic hermit unknown to anyone uniquely held that view, while a famous pre-Socratic philosopher -- called by a name that has come down to us as 'Thales' -- held a different view, which was misunderstood by his contemporaries and transformed in the passage through time into the view that all is water. Then, it seems, the student's use of 'Thales' would refer to the famous philosopher, despite being guided by a description of the hermit, contrary to weak descriptivism.

A different argument involves speakers who can uniquely describe the referent of their use of a name *n*, but only relying on other names the referents of which they can uniquely describe only by using *n*. Consider *x*, whose only knowledge of Cicero is that he was the Roman orator who denounced the senator Catiline, and whose only knowledge of Catiline is that he was the senator denounced by the Roman orator Cicero. If *x* were given the referent of either name, *x* could descriptively fix the referent of the other. But *x* can't do either, since all *x* knows is that some Roman orator denounced Roman senator, which, we may assume, doesn't determine the referent of either name. Nevertheless *x* uses 'Cicero' to designate Cicero and 'Catiline' to designate Catiline. Kripke mentions that some are in similar shape with 'Einstein' and 'Relativity Theory'.

Such circularity isn't unusual. Most descriptions we might volunteer when asked about the reference of names (or natural kind terms) include other names (or natural kind terms), often producing widespread, interconnected dependencies, which, when considered together, don't uniquely determine referents of any of the terms. The challenge to producing *any* unique, *purely qualitative* grounding of names and kind terms is, therefore, daunting. The idea that our entire network of such terms can be given a global reference-determining descriptive grounding is an unsupported article of faith.

Kripke did believe that in rare cases names *are* introduced by stipulating that they are to rigidly designate whatever is denoted by the description that introduces them -- as when one says "Let 'Jack the Ripper' name the person who committed such-and-such crimes." In such cases, Kripke seemed to suggest, *merely understanding n* puts one in a position to know that $\lceil n$ is/was *D* (if anything was) \rceil is true -- where *D* was used in the stipulation. In such cases associating *n* with a particular descriptive content is taken (temporarily) as a common

achievement of all who understand n. But nothing like this is true of ordinary names. For them, there are no privileged (reference-determining) descriptions and no special semantic knowledge apart from that expressed by '[n' designates n]'. With an ordinary name, to understand it is just to use it to refer to the same thing other users do -- no matter how this is accomplished.

Competence with such a name n involves having a referential intention determining something o as referent, and awareness that to use n is to say something of o. Although one can acquire the needed intention simply by speaking a language in which n already names o, and intending to use n as others do, this is a general fact about language applying to many kinds of words. not a special semantic fact about names (or natural kind terms).

The Historical-Chain Model of Reference-Fixing and Transmission

After criticizing strong and weak descriptivism, Kripke offered commonsense platitudes pointing toward a positive theory of reference-fixing. First, he observed, a name is introduced, either by stipulation -- "I name this ship 'The Queen Mary'," or "Let's call him 'Greg'" -- or by a recurring pattern of uses of a new phrase, e.g. 'Green Lake', to refer to some body of water. Once introduced, the name is used in conversation to refer to its bearer. New people pick it up intending to refer to the whatever their sources did. The process continues, producing in a chain of reference transmission. Sometime descriptive content accompanies passing the name from one user to another. But this content needn't be accurate or reference determining. Normally, reference is determined by the chain itself. If one who acquires a name intends to use it to refer to whatever one's sources used it to designate, it may not matter which other beliefs one may have. Instead of those beliefs determining reference, often the chain itself partially determines the contents of one's beliefs. In this way, referring and believing become, in part, community efforts.

Kripke didn't offer this commonsense sketch as a precise or comprehensive theory. For example, no attempt was made to answer Q.

Q What, if anything, does x's use of n designate if (a) x stands in a chain governing n grounded in o, (b) x intends to use n to refer to whatever grounds the chain, (c) x also intends to use n to refer to o* with which x is acquainted or to which x is connected by a significant description, (d) $o \neq o^*$, and (e) x wrongly believes that o* grounds the chain.

The Madagascar example discussed in Evans (1973) and Berger (2002) is a case of this type. Aside from disputed historical details, the basic idea is that newcomers pick up the name 'Madagascar' from locals, intending to use it to preserve its previous reference, while wrongly assuming it had designated the big island off the east coast of Africa (rather than a portion of the continent). Because the island is what the newcomers wished to communicate about, it ultimately became the referent of the name. This suggests that although historical chains of reference transmission are weighty factors in determining the referents of uses of names, the dynamics of communicative situations also play a role. A genuine theory of reference determination, which Kripke didn't offer, would make proposals about this, while marshaling evidence to verify it. There are various ways of coming to have a referential intention regarding a name that may conflict with the intention of preserving its previous reference, thus, making reference assignments not determined by chains of reference transmission possible.

Although Kripke didn't tell us how to determine reference in these cases, he did provide an illuminating example on which one might build. Kripke (1979) discusses a case in which Mary and Janet see Smith in the distance, mistaking him for Jones. Mary says "Look, Jones is raking leaves" thereby saying something true about Smith and something false about Jones. Because of her conflicting intentions, her use of 'Jones' can plausibly be understood as referring to both. If Mary were to immediately add, "Run over and warn him to get inside; a tornado is about to hit," her reference to Smith might be paramount. Depending on how the scenario develops, this

could lead to Smith becoming the temporary, or permanent, referent of 'Jones' in the language of the community.⁵

Although I haven't yet touched on parasitic descriptions that identify the referent of one's use of *n* as being whomever or whatever certain others use *n* to designate, they doesn't change the basic picture. Consider Kripke's 'Peano' example. Most who have heard of Peano believe he discovered the now standard Peano Axioms of arithmetic. But he didn't. Although he published them, he credited Richard Dedekind with the formalization. Since many who use 'Peano' to refer to Peano don't know this, the description they most strongly associate with the name -- *the one who first formalized arithmetic* -- doesn't designate their referent. Nor does *the person to whom most people refer when they use 'Peano'* unless most 'Peano' users have other descriptions that uniquely pick him out, which, we may presume, they don't. Even *the person to whom most experts, i.e., mathematicians, refer when they use 'Peano'* may not suffice. Since mathematics is a big field, there is no guarantee that most mathematicians are aware of the provenance of the Peano Axioms. What about *the person to whom most users of the name attribute such-and-such axioms*? That's no good either, since it requires independently identifying to whom most people refer when they say, "Peano discovered those axioms." Thus, we still have no way of saving weak descriptivism.

The Architecture of Causal Descriptivism about Names

Despite these problems, dedicated descriptivists like David Lewis thought that descriptivism's salvation was lying in plain sight.

⁵ Considerations like these must be born in mind when people speak of different "intuitions" about what uses of names refer to when presented with Kripke-style scenarios. The questions must be put carefully, indicating conversational dynamics. Even Kripke's 'Godel'-'Schmidt' scenario (in which the one baptized 'Godel' steals the work of the one baptized 'Schmidt') can be contextualized to elicit the judgment that a given use of the 'Godel' refers not to the one grounding the 'Godel' chain but to Schmidt. This can happen in a scenario in which what is being discussed is why at a certain stage of the work "Godel" made one argumentative move rather than another. Such cases don't falsify Kripke's platitudinous historical-chain picture.

Did not Kripke and his allies refute the description theory of reference, at least for names of people and places?...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 'Toromeo'' or maybe 'the causal source of this token: Toromeo', and for the account of the relation being invoked here, just consult the writings of causal theorists of reference.⁶

Lewis's thought builds on Kripke's picture of reference transmission in which speakers pick up proper names (and other kinds of words) and use them to designate what others do. Because speakers routinely use language to communicate, they know this. They realize that many of their words have *somehow* inherited their reference from that of others, without knowing precisely how this occurs.⁷ A sophisticated speaker might even associate *n* with the description *D* -- *the referent of those uses of n from which my present use of n somehow inherits its reference*. But *D* doesn't semantically fix the referent of her use of *n*. For *D* to do that, satisfaction of *D* must be the mechanism by which her reference to *o* is established, in which case the fact that *o* satisfies *D* must be *the reason* she refers to *o* when using *n*. It isn't. On the contrary, the pattern of explanation is the opposite. The only reason *D* picks out *o* as her referent is because some other, unidentified, process has already determined that her use of *n* refers to *o*. When circularity-generating descriptions like *D* are avoided, the best we can normally provide are descriptive approximations that pick out the right referent in some but not all cases. Thus parasitic descriptions don't save weak descriptivism.⁸

⁶ P. 332, n22 of Lewis (1999).

⁷ For further discussion see pp. 366-371 of Soames (2003) and pp. 299-302 of Soames (2005).

⁸ Lewis missed this because he failed to see the circularity of his description 'the place I have heard of under the name "Toromeo"'. A similar example is given on p. 210 of Jackson (1998b), which is discussed on pp. 186-188 of Soames (2005).

Is there another way of insuring the existence of descriptions required by weak descriptivism? Recall the argument from Jackson (1998b), summarized in the first paragraph of this paper. He says:

If speakers can say what refers to what when various possible worlds are described to them [in Kripkean thought experiments used to refute proposals about which descriptions fix referents of which names], description theorists can identify the property associated in their minds with, for example, the word 'water' [or with the name 'Thales']: *it is the disjunction of the properties that guide the speakers in each particular possible world when they say which stuff [or which individual], if any, counts as water [or Thales]*. The disjunction is in their minds in the sense that they can deliver the answer for each possible world when it is described to them in sufficient detail, but it is implicit in the sense that the pattern that brings the various disjuncts together as part of the possibly highly complex, disjunction may be one they cannot state.⁹

This argument, which is meant to apply to names and natural kind terms, purports to show that each has its reference fixed by a description semantically associated with it by users, even though no particular descriptions have been shown to do so. Surprisingly, the argument is based on Kripkean demonstrations that familiar candidate descriptions *don't* determine reference. Because, Jackson thinks, we have reliable intuitions about what terms we understand designate in the imaginable scenarios that Kripke relies on in his antidescriptivist demonstrations, we *must* be guided by internalized descriptions determining their referent in all scenarios.

There are three main problems with this argument. First, the ability to refute particular descriptivist analyses by finding, for each proposed description D, at least one world-state w at which the referent of n (as used by speakers in w in the way we use n at @) differs from the denotation of D at w, does *not* presuppose that we have ability, required Jackson's argument, to correctly identify, *for every possible world-state*, what, if anything, n designates there. It is perfectly possible, I think likely, that *some* scenarios involving conflicting referential intentions (of the sort illustrated above) would *not* result in uniform judgments by competent users of the name, even when that the name *does* designate something. It is a familiar fact about theorizing

⁹ P. 212 of Jackson (1998b), my emphasis.

that the best account of clear instances of an important linguistic phenomenon often generalize to initially unclear cases in ways that ordinary competent speakers don't uniformly recognize. Although our theories rest on a base of clear pretheoretic judgments, no defensible methodology limits correct extensions of central linguistic concepts to those that ordinarily competent speakers reliably recognize to be so.

The second problem is the argument's need to aggregate descriptions extractable from individual judgments of *what n, as used in w, refers to* into a single reference-fixing description that specifies what n designates at each world-state (at which it is used). If the master description MD for n is to vindicate weak descriptivism, it must be non-circular in two senses. First, satisfaction of MD, at any w must be the mechanism by which the referent of n at w is established. Second, MD can't contain any other term for which descriptivism is supposed to hold, unless the referent of that term, at any world-state, can be determined entirely independently of that of n. Somehow this must be guaranteed for all names, natural kind terms, and related expressions.

To do this, Jackson must maintain either (a) that for each name, natural kind term, or related expression, we can correctly determine its referent in possible scenarios *described without using any such terms*, or (b) that there are some such expressions for which we can do this, which can then be used in describing world-states in ways that allow us to extract non-circular reference-fixing descriptions for all remaining names, natural kind terms, and related expressions, or (c) that using all our pretheoretic intuitions about what refers to what in different scenarios, we can extract a single statement, *expressible without any names, natural kind terms or related expressions*, that fixes the referents of all such expressions in all worlds simultaneously. It has never been shown that this can be done.

The third problem with Jackson's argument is that he, like other weak descriptivists, mistakes the semantic question *What does a term mean, and what does it designate?* with the presemantic question *How did the term initially come to mean and refer to what it does, and how is that meaning and reference maintained?* If *n* is a name, there will be an initial use, or series of uses, by one or more individuals to pick out a referent *o*, followed by a chain of reference transmission in which subsequent uses inherit the reference and semantic content of earlier uses. What does this have to do with semantics? *That n designates o, which is also its semantic content*, is a semantic fact -- a linguistic convention. For each different name in the language of a linguistic community, there is a different convention of this sort plus interpretive conventions governing other words, phrases and sentences. The semantics of a language is a set of such conventions.

By contrast, the fact that this, or that, individual originally introduced *n* to name *o*, after which *n* was passed down a historical chain to others without changing reference is a presemantic fact about how *n* came, and continues, to have the reference it does. It is neither a convention nor something present speakers need to know. Since they are familiar with how language is used to communicate, they will take it for granted that uses of names and many other words often inherit their reference from earlier uses. This is a general fact about linguistic communication, not a semantic fact about their language.¹⁰

The failure of causal descriptivists to vindicate weak descriptivism wasn't their only problem. Taking themselves to have succeeded in vindicating descriptive reference-fixing, Jackson and Lewis needed to rigidify their descriptions to match the modal profile of names, and to bring the epistemic profile of sentences containing their descriptions with those

¹⁰ This point is discussed more fully in response to Jackson (1998a) on pp. 182-184 of Soames (2005).

containing names. As suggested above, and argued in in detail in chapter 2 of Soames (2002) and chapters 5-10 of Soames (2005), this strategy fails at every step.

Natural Kind Terms

Next a word extending the anti-descriptivist account of names to natural kind terms like 'water', 'tiger', 'green', and 'heat', which apply to various objects, or quantities of stuff. Depending on their grammatical categories, these terms, or their close relatives, combine with the copula to form predicates -- 'is water', 'is a tiger', and 'is hot'. In the case of the latter, there is also the comparative form 'is hotter than' which is true of a pair the first of which contains more heat (molecular motion) than the second. Each of the four general terms designates a kind. What is the difference between natural kind terms and other terms? Though Kripke says little about this, an answer can be abstracted from his examples.

Like names, natural kind terms are not synonymous with descriptions associated with them by speakers. Their introduction and transmission are also similar to that of names. Just as ordinary names are often introduced by stipulating that they are to refer to certain individuals with which one is acquainted, so natural kind terms can be introduced to designate kinds with which one is acquainted through their instances. For example, we may imagine 'water' and 'green' being introduced by the following stipulations:

Let 'water' designate the property possessed by (nearly) all members of a certain class of samples that explains their most salient characteristics – e.g., the fact that they boil and freeze at certain temperatures, that they are clear, potable, and necessary to life, etc. The predicate 'is water' applies (at a world-state) to precisely those quantities that have the physical constitution which, at @, explains the salient features of (nearly) all our samples.

Let 'green' designate the property possessed by (nearly) all members of a certain class of sample object surfaces that is causally responsible for the fact that they appear visually similar to us (and different from other surfaces). Thus, 'is green' applies (at a world-state) to all and only those objects the surfaces of which have the characteristic that, at @, causally explains why the our samples look similar to us (and different from other samples).

These stipulations are, of course, idealized. 'Water' wasn't explicitly introduced this way, but it could have been, and it behaves pretty much as if it had been. Presumably, speakers simply started calling certain quantities 'water', intending it to apply, not only to samples they had encountered, but to all other instances "of the same kind." They noticed important properties $P_1 \dots P_n$ of nearly all quantities they had been calling 'water', which they assumed to have a common unifying explanation, which, they thought, made them instances of the kind of stuff they are. Hoping to track this common element, they used 'water' to designate the property (which they then had no more informative way of designating) possession of which explained the fact that nearly all their samples had $P_1 \dots P_n$. Because their explanatory assumption turned out to be correct, the payoff was a term that can be used in counterfactual explanations and law-like generalizations in ways that a mere conjunction of terms designating $P_1 \dots P_n$ can't.

Other general terms -- e.g. 'tiger', 'gold', 'heat', and (by extension) 'hot' -- fit this picture. So does 'green', which designates a family of specific shades. The samples associated with the word are objects each of which is an instance of one of the shades. What the samples have in common is that they are judged by humans with normal color vision to look similar to one another (in various contexts and lights), and different from objects of other shades outside the family (in similar contexts and lights). 'Green' designates the property of object surfaces that, at @, explains the perceived differences or similarities.

So far, I have spoken only of how natural kind terms are introduced. Once introduced, they are typically passed from speaker to speaker in a chain of reference transmission, just as names are. Hence, it's not surprising that arguments against descriptive analyses of the meaning or reference of natural kind terms largely parallel those against descriptive analyses of names. The most telling arguments against weak (and thereby against strong) descriptivism for natural

kind terms are based on our ability to use them correctly despite the ignorance, error, and incompleteness of our descriptions of them. Most of us routinely use many natural kind terms that we can't uniquely and non-circularly describe. In my case these include 'bauxite', 'tungsten' 'obsidian', 'wildebeest', 'sycamore', 'pulsar', 'insect', and many more. Even when we can describe a natural kind, as (perhaps) in (2), our descriptions aren't synonyms for the kind terms.

2. *Light is the cause of visual experience.*

Heat is molecular motion.

Humans are primates that are not apes, monkeys or lemurs.

Carbon dioxide is the chemical element extracted from the air and used by plants in photosynthesis.

Moreover, all though some of the propositions in (2) are necessary, none are apriori. In short, most of Kripke's arguments against descriptive analyses of proper names, carry over to similar analyses of natural kind terms.

Saving Important Insights of Causal Descriptivism

Kripke was right that descriptions don't give the meanings, or semantically fix the referents, of most ordinary names and natural kind terms. Post-Kripkean nondescriptivists were right that their semantic contents are the objects or kinds they designate. Causal descriptivists were right that there is more to the meanings of these terms than those objects or kinds. If that sounds incoherent, it is probably because you are thinking, along with leading proponents on both sides of the debate, that meaning, which either is or determines reference, and semantic content are the same. They aren't. What must be recognized, but in most quarters hasn't been, is that meaning, in the sense of what is understood and widely presupposed by competent language users, can affect the contents of assertions, and the beliefs expressed by them, without affecting semantic content.¹¹

¹¹ See chapter 4 of Soames (2015) and also my "Philosophy of Language in the 21st Century," forthcoming.

The semantic content of a term is what it contributes to the compositionally determined semantic contents of all sentential clauses in which it occurs, including those governed by modal operators (e.g. *necessarily, possibly*). This, in the case of names and natural kind terms, is the object or kind the term designates. But there is more to understanding certain special names and many ordinary natural kind terms than simply being able to use them to designate their semantic contents. For example, one who uses ‘Hesperus’ and ‘Phosphorus’ is expected know that users typically presuppose that ‘Hesperus’ stands for something visible in the evening and ‘Phosphorus’ stands for something visible in the morning. One who mixes this up *misunderstands* the names. Those who know enough to use the names are aware of this.

With this in mind, suppose that A assertively utters (3) addressing B in a context in which they commonly presuppose that each understands the names.

3. Hesperus is Phosphorus

Although A thereby asserts the trivial proposition that Venus is Venus which is the semantic content of (3), B extracts more information from A’s utterance. Presupposing that A *understands* the names, B reasons that A knows she will be taken to be committed to the claim that *the unique object that is both Hesperus and visible in the evening is the unique object that is both Phosphorus and visible in the morning*. Knowing that A expects B to see this, he correctly concludes that A *asserted* the descriptively enriched proposition.

Although this proposition is contingent, A’s use of (4) asserts something true, and nothing false.

4. It is necessarily true that Hesperus is Phosphorus.

The reason no falsehood is asserted is that what understanding the names requires is *knowing that most agents who use them take, and expect others to take, ‘Hesperus’ to stand for*

something (actually) seen in the evening and 'Phosphorus' to stand for something (actually) seen in the morning. Since taking the names to refer to things *actually* seen at certain times tells one *nothing* about when they are seen at other *possible* world-states, A and B *don't* descriptively enrich the occurrences of the names under the modal operator in (4).

A similar account applies to (5).

5. Water is H₂O.

We may imagine that one who understands (5) uses 'is' to stand for identity, 'water' to designate a natural kind *k*, and the name 'H₂O' (which is related to, but semantically distinct from, the phrase 'the substance molecules of which consist of two hydrogen atoms and one oxygen atom') to rigidly designate *k*. Since 'water' and 'H₂O' have the same semantic content, the semantic content of (5) is the triviality *that k = k*, which isn't what people intend to communicate. In most cases in which (5) is used, speaker-hearers mutually presuppose that they *understand* the terms. In such cases, a speaker A asserts the proposition *that k = k*, plus another proposition. Presupposing that A *understands* the terms, B reasons that A knows that she will be taken to be committed to the claim that *k* is both *a chemical compound involving hydrogen and oxygen* (which is what understanding 'H₂O' requires) *and one instances of which are clear and potable, necessary for life, and found in lakes and rivers* (which understanding 'water' requires). Realizing that A expects him to so reason, B correctly concludes that A *asserted* this informative, descriptively enriched proposition. The explanation of why, nevertheless, *Necessarily water is H₂O* is true parallels the explanation of why *Necessarily Hesperus is Phosphorus* is true.

The general lesson here is that our pretheoretic conception of *meaning* incorporates both elements of what is ordinarily called *understanding* and what theorists call *semantic content*.

Understanding is a graded term. One can understand an expression or sentence more or less well. In addition to coming in degrees, understanding is also context sensitive; e.g., what counts as (adequately) understanding technical terms depends on the group or setting in which the language is used. All of this points to the need for richer and more integrated semantic-pragmatic theories of the contents of illocutionary acts than were common in the heyday of disputes over causal-descriptivism. If, as seems undeniable, asserted content arises from *semantic contents* plus contents of widespread presuppositions associated with *understanding*, then a more nuanced distinction between semantic content and *illocutionary content* is needed.¹²

This is only the beginning of a much longer story incorporating new analyses incorporating representationally identical but cognitively distinct propositions distinguished by the presence or absence of fine-grained cognitive, non-Fregean, modes of presentation, new accounts of propositional attitudes sensitive to such differences, and new accounts of the relationship between sentences and the multiple propositions that single, unambiguous sentences may simultaneously express, and that single utterances of them may assert.¹³ This longer story is needed, in part, because the semantic-pragmatic frameworks in which the dispute between descriptivist's and antidescriptivists took place were too impoverished to accommodate the respects in which each side was right, in which each went wrong.

¹² See chapter 7 of Soames (2010) for some tentative first steps.

¹³ For an introduction, see Soames forthcoming; for more details see chapters 2 - 8 of Soames (2015).

References

(*Suggestions for Further Reading)

*Berger, Alan (2002) *Terms and Truth*. Cambridge MA: MIT Press.

Jackson, Frank (1998a) *From Metaphysics to Ethics*. Oxford: Oxford University Press.

*_____, (1998b) "Reference and Description Revisited," *Philosophical Perspectives* 12: 201-18.

Kaplan, David (1989) "Demonstratives," in Joseph Almog, John Perry, and Howard Wettstein, eds. *Themes from Kaplan*. New York: Oxford University Press.

*Kripke, Saul (1979) "Speaker's Reference and Semantic Reference," in Peter French, Theodore Uehling, and Howard Wettstein, eds. *Contemporary Perspectives in the Philosophy of Language*. Minneapolis: University of Minnesota.

_____, (1980) *Naming and Necessity*, Cambridge MA.: Harvard University Press.

Lewis, David (1999) "Naming the Colors." *Papers in Metaphysics and Epistemology*. Cambridge. Cambridge University Press.

Soames, Scott (2002) *Beyond Rigidity*. New York: Oxford University Press.

_____, (2003) *Philosophical Analysis in the Twentieth Century*, Vol. 2. Princeton and Oxford: Princeton University Press.

_____, (2005) *Reference and Description*. Princeton and Oxford. Princeton University Press.

_____, (2010) *Philosophy of Language*, Princeton and Oxford. Princeton University Press.

*_____, (2015) *Rethinking Language, Mind, and Meaning*. Princeton and Oxford. Princeton University Press.

*_____, forthcoming. "Philosophy of Language in the Twenty-first Century" in *The Cambridge Handbook of Philosophy of Language*, Piotr Stalmaszczyk (ed.).