

## The Causal Theory of Names

1. In a paper which provides the starting-point of this enquiry Saul Kripke opposes what he calls the Description Theory of Names and makes a counter-proposal of what I shall call the Causal Theory.<sup>1</sup> To be clear about what is at stake and what should be the outcome in the debate he initiated seems to me important for our understanding of talk and thought about the world in general as well as for our understanding of the functioning of proper names. I am anxious therefore that we identify the profound bases and likely generalizations of the opposing positions and do not content ourselves with counter-examples.

I should say that Kripke deliberately held back from presenting his ideas as a theory. I shall have to tighten them up, and I may suggest perhaps unintended directions of generalization; therefore his paper should be checked before the Causal Theory I consider is attributed to him.

There are two related but distinguishable questions concerning proper names. The first is about what the name denotes upon a particular occasion of its use when this is understood as being partly determinative of what the speaker strictly and literally said. I shall use the faintly barbarous coinage: *what the speaker denotes* (upon an occasion) for this notion. The second is about *what the name denotes*; we want to know what conditions have to be satisfied by an expression and an item for the first to be the, or a, name of the second. There is an entirely parallel pair of questions concerning general terms. In both cases it is ambiguity

From *ASSV* 47 (1973), pp. 187-208. Reprinted by permission of the Aristotelian Society.

<sup>1</sup> S. A. Kripke, 'Naming and Necessity', in D. Davidson and G. Harman (eds). *Semantics of Natural Language* (Dordrecht: Reidel, 1972), pp. 253-355 (+ Appendix).

which prevents an easy answer of the first in terms of the second; to denote  $x$  it is not sufficient merely to utter something which is  $x$ 's name.

Consequently there are two Description Theories, not distinguished by Kripke.<sup>2</sup> The Description Theory of speaker's denotation holds that a name 'NN' denotes  $x$  upon a particular occasion of its use by a speaker  $S$  just in case  $x$  is uniquely that which satisfies all or most of the descriptions  $\phi$  such that  $S$  would assent to 'NN is  $\phi$ ' (or 'That NN is  $\phi$ '). Crudely: the cluster of information  $S$  has associated with the name determines its denotation upon a particular occasion by *fit*. If the speaker has no individuating information he will denote nothing.

The Description Theory of what a name denotes holds that, associated with each name as used by a group of speakers who believe and intend that they are using the name with the same denotation, is a description or set of descriptions cullable from their beliefs which an item has to satisfy to be the bearer of the name. This description is used to explain the role of the name in existential, identity, and opaque contexts. The theory is by no means committed to the thesis that every user of the name must be in possession of the description; just as Kripke is not committed to holding that every user of the expression 'one metre' knows about the metre rod in Paris by saying that its reference is fixed by the description 'Length of stick  $S$  in Paris'. Indeed if the description is arrived at in the manner of Strawson<sup>3</sup>—averaging out the different beliefs of different speakers—it is most unlikely that the description will figure in every user's name-associated cluster.

The direct attack in Kripke's paper passes this latter theory by; most conspicuously the charge that the Description Theory ignores the social character of naming. I shall not discuss it explicitly either, though it will surface from time to time and the extent to which it is right should be clear by the end of the paper.

<sup>2</sup> This can be seen in the way the list of theses defining the Description Theory alternate between those mentioning a speaker and those that don't, culminating in the uneasy idea of an idiolect of one. The Description Theorists of course do not themselves distinguish them clearly either, and many espouse both.

<sup>3</sup> P. F. Strawson, *Individuals* (London: Methuen, 1959), p. 191.

Kripke's direct attacks are unquestionably against the first Description Theory. He argues:

(a) An ordinary man in the street can denote the physicist Feynman by using the name 'Feynman' and say something true or false of him even though there is no description uniquely true of the physicist which he can fashion. (The conditions aren't necessary.)

(b) A person who associated with the name 'Gödel' merely the description 'prover of the incompleteness of Arithmetic' would none the less be denoting Gödel and saying something false of him in uttering 'Gödel proved the incompleteness of Arithmetic' even if an unknown Viennese by the name of Schmidt had in fact constructed the proof which Gödel had subsequently broadcast as his own. (If it is agreed that the speaker does not denote Schmidt the conditions aren't sufficient; if it is also agreed that he denotes Gödel, again they are not necessary.)

The strong thesis (that the Description Theorist's conditions are sufficient) is outrageous. What the speaker denotes in the sense we are concerned with is connected with saying in that strict sense which logicians so rightly prize, and the theory's deliverance of strict truth conditions are quite unacceptable. They would have the consequence, for example, that if I was previously innocent of knowledge or belief regarding Mr  $Y$ , and  $X$  is wrongly introduced to me as Mr  $Y$ , then I must speak the truth in uttering 'Mr  $Y$  is here' since  $X$  satisfies the overwhelming majority of descriptions I would associate with the name and  $X$  is there. I have grave doubts as to whether anyone has ever seriously held this thesis.

It is the weaker thesis—that some descriptive identification is necessary for a speaker to denote something—that it is important to understand. Strictly, Kripke's examples do not show it to be false since he nowhere provides a convincing reason for not taking into account speakers' possession of descriptions like 'man bearing such-and-such a name'; but I too think it is false. It can be seen as the fusion of two thoughts. First: that in order to be saying something by uttering an expression one must utter the sentence with

certain intentions; this is felt to require, in the case of sentences containing names, that one be aiming at something with one's use of the name. Secondly—and this is where the underpinning from a certain Philosophy of Mind becomes apparent—to have an intention or belief concerning some item (which one is not in a position to demonstratively identify) one must be in possession of a description uniquely true of it. Both strands deserve at least momentary scrutiny.

We are prone to pass too quickly from the observation that neither parrots nor the wind *say* things to the conclusion that to say that  $p$  requires that one must intend to say that  $p$  and therefore, so to speak, be able to identify  $p$  independently of one's sentence. But the most we are entitled to conclude is that to say something one must intend to say something by uttering one's sentence (one normally will intend to say what it says). The application of the stricter requirement would lead us to relegate too much of our discourse to the status of mere mouthing. We constantly use general terms of whose satisfaction conditions we have but the dimmest idea. 'Microbiologist', 'chlorine' (the stuff in swimming-pools), 'nicotine' (the stuff in cigarettes); these, and countless other words, we cannot define nor offer remarks which would distinguish their meaning from that of closely related words. It is wrong to say that we say nothing by uttering sentences containing these expressions, even if we recoil from the strong thesis, from saying that what we do say is determined by those hazy ideas and half-identifications we would offer if pressed.

The Philosophy of Mind is curiously popular but rarely made perfectly explicit.<sup>4</sup> It is held by anyone who holds that  $S$  believes that  $a$  is  $F$  if and only if

$$\exists \phi [S \text{ believes } \exists x(\phi x \ \& \ (\forall y)(\phi y \rightarrow x = y) \ \& \ Fx) \ \& \ \phi a \ \& \ (\forall y)(\phi y \rightarrow y = a)]$$

<sup>4</sup> See, e.g., J. R. Searle, *Speech Acts* (Cambridge: Cambridge University Press, 1969), p. 87; E. Gellner, 'Ethics and Logic', *Proceedings of the Aristotelian Society* 55 (1954-5), pp. 157-78; B. Russell, *Problems of Philosophy* (Oxford: Oxford University Press, paperback 1976), p. 29. E. Sosa criticizes it in 'Quantifiers Belief and Sellars', in J. W. Davis, D. J. Hockney, and W. K. Wilson (eds), *Philosophical Logic* (Dordrecht: Reidel, 1969), p. 69.

Obvious alterations would accommodate the other psychological attitudes. The range of the property quantifier must be restricted to exclude such properties as 'being identical with  $a$ ', otherwise the criterion is trivial.<sup>5</sup> The situation in which a thinking, planning or wanting human has some item which is the object of his thought, plan or desire is represented as a species of essentially the same situation as that which holds when there is no object and the thought, plan or desire is, as we might say, purely general. There are thoughts, such as the thought that there are eleven-fingered men, for whose expression general terms of the language suffice. The idea is that when the psychological state involves an object, a general term believed to be uniquely instantiated and in fact uniquely instantiated by the item which is the object of the state will figure in its specification. This idea may be coupled with a concession that there are certain privileged objects to which one may be more directly related; indeed such a concession appears to be needed if the theory is to be able to allow what appears an evident possibility: object-directed thoughts in a perfectly symmetrical or cyclical universe.

This idea about the nature of object-directed psychological attitudes obviously owes much to the feeling that there must be something we can say about what is believed or wanted even when there is no appropriate object actually to be found in the world. But it can also be seen as deriving support from a Principle of Charity: so attribute objects to beliefs that true belief is maximized. (I do not think this is an acceptable principle; the acceptable principle enjoins minimizing the attribution of *inexplicable* error and therefore cannot be operated without a theory of the causation of belief for the creatures under investigation.)

We cannot deal comprehensively with this Philosophy of Mind here. My objections to it are essentially those of Wittgenstein. For an item to be the object of some psychological attitude of yours may be simply for you to be placed in a context which relates you to that thing. What makes it one rather than the other of a pair of identical twins that you are in love with? Certainly not some specification blueprinted

<sup>5</sup> I owe this observation to G. Harman.

in your mind; it may be no more than this: it was one of them and not the other that you have met. The theorist may gesture to the description 'the one I have met' but can give no explanation for the impossibility of its being outweighed by other descriptions which may have been acquired as a result of error and which may in fact happen to fit the other, unmet, twin. If God had looked into your mind, he would not have seen there with whom you were in love, and of whom you were thinking.

With that I propose to begin considering the Causal Theory.

2. The Causal Theory as stated by Kripke goes something like this. A speaker, using a name 'NN' on a particular occasion will denote some item  $x$  if there is a causal chain of *reference-preserving links* leading back from his use on that occasion ultimately to the item  $x$  itself being involved in a name-acquiring transaction such as an explicit dubbing or the more gradual process whereby nicknames stick. I mention the notion of a reference-preserving link to incorporate a condition that Kripke lays down; a speaker  $S$ 's transmission of a name 'NN' to a speaker  $S'$  constitutes a reference-preserving link only if  $S$  intends to be using the name with the same denotation as he from whom he in his turn learned the name.

Let us begin by considering the theory in answer to our question about speaker's denotation (i.e., at the level of the individual speaker). In particular, let us consider the thesis that it is *sufficient* for someone to denote  $x$  on a particular occasion with the name that this use of the name on that occasion be a causal consequence of his exposure to other speakers using the expression to denote  $x$ .

An example which might favourably dispose one towards the theory is this. A group of people are having a conversation in a pub, about a certain Louis of whom  $S$  has never heard before.  $S$  becomes interested and asks: 'What did Louis do then?' There seems to be no question but that  $S$  denotes a particular man and asks about him. Or on some subsequent occasion  $S$  may use the name to offer some new thought to one of the participants: 'Louis was quite right to do that.' Again he clearly denotes whoever was the subject of conversation in the pub. This is difficult to reconcile with

the Description Theory since the scraps of information which he picked up during the conversation might involve some distortion and fit someone else much better. Of course he has the description 'the man they were talking about' but the theory has no explanation for the impossibility of its being outweighed.

The Causal Theory can secure the right answer in such a case but I think deeper reflection will reveal that it too involves a refusal to recognize the insight about contextual determination I mentioned earlier. For the theory has the following consequence: that at any future time, no matter how remote or forgotten the conversation, no matter how alien the subject-matter and confused the speaker,  $S$  will denote one particular Frenchman—perhaps Louis XIII—so long as there is a causal connection between his use at that time and the long-distant conversation.

It is important in testing your intuitions against the theory that you imagine the predicate changed—so that he says something like 'Louis was a basketball player' which was not heard in the conversation and which arises as a result of some confusion. This is to prevent the operation of what I call the 'mouthpiece syndrome' by which we attach sense and reference to a man's remarks only because we hear someone else speaking through him; as we might with a messenger, carrying a message about matters of which he was entirely ignorant.

Now there is no knock-down argument to show this consequence unacceptable; with pliant enough intuitions you can swallow anything in philosophy. But notice how little *point* there is in saying that he denotes one French king rather than any other, or any other person named by the name. There is now nothing that the speaker is prepared to say or do which relates him differentially to that one king. This is why it is so outrageous to say that he believes that Louis XIII is a basketball player. The notion of saying has simply been severed from all the connections that made it of interest. Certainly we did not think we were letting ourselves in for this when we took the point about the conversation in the pub. What has gone wrong?<sup>6</sup>

<sup>6</sup> Kripke expresses doubts about the sufficiency of the conditions for this sort of reason, see op. cit., p. 303.

The Causal Theory again ignores the importance of surrounding context, and regards the capacity to denote something as a magic trick which has somehow been passed on, and once passed on cannot be lost. We should rather say: in virtue of the context in which the man found himself the man's dispositions were bent towards one particular man—Louis XIII—whose states and doings alone he would count as serving to verify remarks made in that context using the name. And of course that context can persist, for the conversation can itself be adverted to subsequently. But it can also disappear so that the speaker is simply not sensitive to the outcome of any investigations regarding the truth of what he is said to have said. And at this point saying becomes detached, and uninteresting.

(It is worth observing how ambivalent Kripke is on the relation between denoting and believing; when the connection favours him he uses it; we are reminded for example that the ordinary man has a false belief about Gödel and not a true belief about Schmidt. But it is obvious that the results of the 'who are they believing about?' criterion are bound to come dramatically apart from the results of the 'who is the original bearer of the name?' criterion, if for no other reason than that the former must be constructed to give results in cases where there is no name and where the latter cannot apply. When this happens we are sternly reminded that 'X refers' and 'X says' are being used in *technical* senses.<sup>7</sup> But there are limits. One could regard the aim of this paper to restore the connection which must exist between strict truth conditions and the beliefs and interests of the users of the sentences if the technical notion of strict truth conditions is to be of interest to us.)

Reflection upon the conversation in the pub appeared to provide one reason for being favourably disposed towards the Causal Theory. There is another connected reason we ought to examine briefly. It might appear that the Causal Theory provides the basis for a general non-intentional answer to the Problem of Ambiguity. The problem is clear enough: What conditions have to be satisfied for a speaker to have said that *p* when he utters a sentence which may

<sup>7</sup> *Ibid.*, p. 348 fn.

appropriately be used to say that *q* and that *r* and that *s* in addition? Two obvious alternative answers are:

- (a) the extent to which it is reasonable for his audience to conclude that he was saying that *p*
- (b) his intending to say that *p*

and neither is without its difficulties. We can therefore imagine someone is hoping for a natural extension of the Causal Theory to general terms which would enable him to explain for example how a child who did not have determinative intentions because of the technical nature of the subject-matter may still say something determinate using a sentence which is in fact ambiguous.

I touch upon this to ensure that we are keeping the range of relevant considerations to be brought to bear upon the debate as wide as it must be. But I think little general advantage can accrue to the Causal Theory from thus broadening the considerations. The reason is that it simply fails to have the generality of the other two theories; it has no obvious application, for example, to syntactic ambiguity or to ambiguity produced by attempts to refer with non-unique descriptions, or pronouns. It seems inconceivable that the general theory of disambiguation required for such cases would be inadequate to deal with the phenomenon of shared names and would require *ad hoc* supplementation from the Causal Theory.

I want to stress how, precisely because the Causal Theory ignores the way context can be determinative of what gets *said*, it has quite unacceptable consequences. Suppose, for example, on a TV quiz programme I am asked to name a capital city and I say 'Kingston is the capital of Jamaica'; I should want to say that I had said something strictly and literally true even though it turns out that the man from whom I had picked up this scrap of information was actually referring to Kingston upon Thames and making a racist observation.

It may begin to appear that what gets said is going to be determined by what name is used, what items bear the name, and general principles of contextual disambiguation. The

causal origin of the speaker's familiarity with the name, save in certain specialized 'mouthpiece cases', does not seem to have a critical role to play.

This impression may be strengthened by the observation that a causal connection between my use of the name and use by others (whether or not leading back ultimately to the item itself) is simply not necessary for me to use the name to say something. Amongst the Wagera Indians, for example, 'newly born children receive the names of deceased members of their family according to strict rules . . . the first born takes on the name of the paternal grandfather, the second that of the father's eldest brother, the third that of the maternal grandfather'.<sup>8</sup> In these and other situations (names for streets in US cities etc.) a knowledgeable speaker may excogitate a name and use it to denote some item which bears it without any causal connection whatever with the use by others of that name.

These points might be conceded by Kripke while maintaining the general position that the denotation of a name in a community is still to be found by tracing a causal chain of reference preserving links back to some item. It is to this theory that I now turn.

3. Suppose a parallel theory were offered to explain the sense of general terms (not just terms for natural kinds). One would reply as follows:

'There aren't two fundamentally different mechanisms involved in a word's having a meaning: one bringing it about that a word acquires a meaning, and the other—a causal mechanism—which operates to ensure that its meaning is preserved. The former processes are operative all the time; whatever explains how a word gets its meaning also explains how it preserves it, if preserved it is. Indeed such a theory could not account for the phenomenon of a word's changing its meaning. It is perfectly possible for this to happen without anyone's intending to initiate a new practice with the word; the causal chain would then lead back too far.'

Change of meaning would be decisive against such a theory

<sup>8</sup> E. Delhaise, 'Les Wagera', *Monogr. Ethnogr.* (1909).

of the meaning of general terms. Change of denotation is similarly decisive against the Causal Theory of Names. Not only are changes of denotation imaginable, but it appears that they actually occur. We learn from Isaac Taylor's *Names and their History* (1898):

In the case of 'Madagascar' a hearsay report of Malay or Arab sailors misunderstood by Marco Polo . . . has had the effect of transferring a corrupt form of the name of a portion of the African mainland to the great African Island.

A simple imaginary case would be this: 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.

It is clear that the Causal Theory unamended is not adequate. It looks as though, once again, the intentions of the speakers to use the name to refer to something must be allowed to count in determination of what it denotes.

But it is not enough to say that and leave matters there. We must at least sketch a theory which will enable 'Madagascar' to be the name of the island yet which will not have the consequence that 'Gödel' would become a name of Schmidt in the situation envisaged by Kripke, nor 'Goliath' a name of the Philistine killed by David. (Biblical scholars now suggest that David did not kill Goliath, and that the attribution of the slaying to Elhannan the Bethlehemite in 2 Sam. 21: 19 is correct. David is thought to have killed a Philistine but not Goliath.)<sup>9</sup> For although this has never been explicitly argued I would agree that even if the 'information' connected with the name in possession of an entire community was merely that 'Goliath was the Philistine David slew' this would still not mean that 'Goliath' referred in that community to that man, and therefore that the sentence expressed a truth. And if we simultaneously thought that the name *would* denote the Philistine slain by Elhannan then both the necessity and sufficiency of the conditions suggested by the Description

<sup>9</sup> H. W. Robinson, *The History of Israel* (London: Duckworth, 1941). p. 187.

Theory of the denotation of a name are rejected. This is the case Kripke should have argued but didn't.

4. Before going on to sketch such a theory in the second part of this paper let me survey the position arrived at and use it to make a summary statement of the position I wish to adopt.

We can see the undifferentiated Description Theory as the expression of two thoughts.

(a) The denotation of a name is determined by what speakers intend to refer to by using the name.

(b) The object a speaker intends to refer to by his use of a name is that which satisfies or fits the majority of descriptions which make up the cluster of information which the speaker has associated with the name.

We have seen great difficulties with (a) when this is interpreted as a thesis at the micro-level. But consideration of the phenomenon of a name's getting a denotation, or changing it, suggests that there being a community of speakers using the name with such-and-such as the intended referent is likely to be a crucial constituent in these processes. With names as with other expressions in the language, what they signify depends upon what we use them to signify; a truth whose recognition is compatible with denying the collapse of saying into meaning at the level of the individual speaker.

It is in (b) that the real weakness lies: the bad old Philosophy of Mind which we momentarily uncovered. Not so much in the idea that the intended referent is determined in a more or less complicated way by the associated information, but the specific form the determination was supposed to take: *fit*. There is something absurd in supposing that the intended referent of some perfectly ordinary use of a name by a speaker could be some item utterly isolated (causally) from the user's community and culture simply in virtue of the fact that it fits better than anything else the cluster of descriptions he associates with the name. I would agree with Kripke in thinking that the absurdity resides in the absence

of the causal relation between the item concerned and the speaker. But it seems to me that he has mislocated the causal relation; the important causal relation lies between that item's states and doings and the speaker's body of information—not between the item's being dubbed with a name and the speaker's contemporary use of it.

Philosophers have come increasingly to realize that major concepts in epistemology and the philosophy of mind have causality embedded within them. Seeing and knowing are both good examples.

The absurdity in supposing that the denotation of our contemporary use of the name 'Aristotle' could be some unknown (n.b.) item whose doings are causally isolated from our body of information is strictly parallel to the absurdity in supposing that one might be seeing something one has no causal contact with solely upon the ground that there is a splendid match between object and visual impression.

There probably is some *degree of fit* requirement in the case of seeing which means that after some amount of distortion or fancy we can no longer maintain that the causally operative item was still being seen. And I think it is likely that there is a parallel requirement for referring. We learn, for example, from E. K. Chambers's *Arthur of Britain* that Arthur had a son Anir 'whom legend has perhaps confused with his burial place'. If Kripke's notion of reference fixing is such that those who said Anir was a burial place of Arthur might be denoting a person it seems that it has little to commend it, and is certainly not justified by the criticism he makes against the Description Theory. But the existence or nature of this 'degree of fit' requirement will not be something I shall be concerned with here.

We must allow, then, that the denotation of a name in the community will depend in a complicated way upon what those who use the term intend to refer to, but we will so understand 'intended referent' that typically a *necessary* (but not sufficient) condition for *x*'s being the intended referent of *S*'s use of a name is that *x* should be the source of causal origin of the body of information that *S* has associated with the name.

## II

5. The aim I have set myself, then, is modest; it is not to present a complete theory of the denotation of names. Without presenting a general theory to solve the problem of ambiguity I cannot present a theory of speaker's denotation, although I will make remarks which prejudice that issue. I propose merely to sketch an account of what makes an expression into a name for something that will allow names to change their denotations.

The enterprise is more modest yet for I propose to help myself to an undefined notion of speaker's reference by borrowing from the theory of communication. But a word of explanation.

A speaker may have succeeded in *getting it across* or in *communicating* that  $p$  even though he uses a sentence which may not appropriately be used to say that  $p$ . Presumably this success consists in his audience's having formed a belief about him. This need not be the belief that the speaker intended to say in the strict sense that  $p$ , since the speaker may succeed in getting something across despite using a sentence which he is known to know cannot appropriately be used to say that  $p$ . The speaker will have referred to  $a$ , in the sense that I am helping myself to, only if he has succeeded in getting it across that  $Fa$  (for some substitution  $F$ ). Further stringent conditions are required. Clearly this notion is quite different from the notion of denotation which I have been using, tied as denotation is to saying in the strict sense. One may refer to  $x$  by using a description that  $x$  does not satisfy; one may not thus denote  $x$ .

Now a speaker may know or believe that there is such-and-such an item in the world and intend to refer to it. And this is where the suggestion made earlier must be brought to bear, for *that* item is not (in general) the satisfier of the body of information the possession by the speaker of which makes it true that he knows of the existence of the item; it is rather that item which is causally responsible for the speaker's possession of that body of information, or dominantly responsible if there is more than one. (The point is of course not specific to this intention, or to intention as

opposed to other psychological attitudes.) Let us then, very briefly, explore these two ideas: source and dominance.

Usually our knowledge or belief about particular items is derived from information-gathering transactions, involving a causal interaction with some item or other, conducted ourselves or is derived, maybe through a long chain, from the transactions of others. Perception of the item is the main but by no means the only way an item can impress itself on us; for example, a man can be the source of things we discover by rifling through his suitcase or by reading his works.

A causal relation is of course not sufficient; but we may borrow from the theory of knowledge and say something like this.  $X$  is the source of the belief  $S$  expresses by uttering ' $Fa$ ' if there was an episode which caused  $S$ 's belief in which  $X$  and  $S$  were causally related in a type of situation apt for producing knowledge that something  $F$ -s ( $\exists x(Fx)$ )—a type of situation in which the belief that something  $F$ -s would be caused by something's  $F$ -ing. That it is a way of producing knowledge does not mean that it cannot go wrong; that is why  $X$ , by smoking French cigarettes, can be the source of the belief  $S$  expresses by ' $a$  smokes Greek cigarettes'.

Of course some of our information about the world is not so based; we may deduce that there is a tallest man in the world and deduce that he is over 6 feet tall. No man is the source of this information; a name introduced in relation to it might function very much as the unamended Description Theory suggested.

Legend and fancy can create new characters, or add bodies of sourceless material to other dossiers; restrictions on the causal relation would prevent the inventors of the legends turning out to be the sources of the beliefs their legends gave rise to. Someone other than the  $\phi$  can be the source of the belief  $S$  expresses by ' $a$  is the  $\phi$ '; Kripke's Gödel, by claiming the proof, was the source of the belief people manifested by saying 'Gödel proved the incompleteness of Arithmetic', not Schmidt.

Misidentification can bring it about that the item which is the source of the information is different from the item about which the information is believed. I may form the belief about the wife of some colleague that she has nice legs



upon the basis of seeing someone else—but the girl I saw is the source.

Consequently a cluster or dossier of information can be dominantly of<sup>10</sup> an item though it contains elements whose source is different. And we surely want to allow that persistent misidentification can bring it about that a cluster is dominantly of some item other than that it was dominantly of originally.

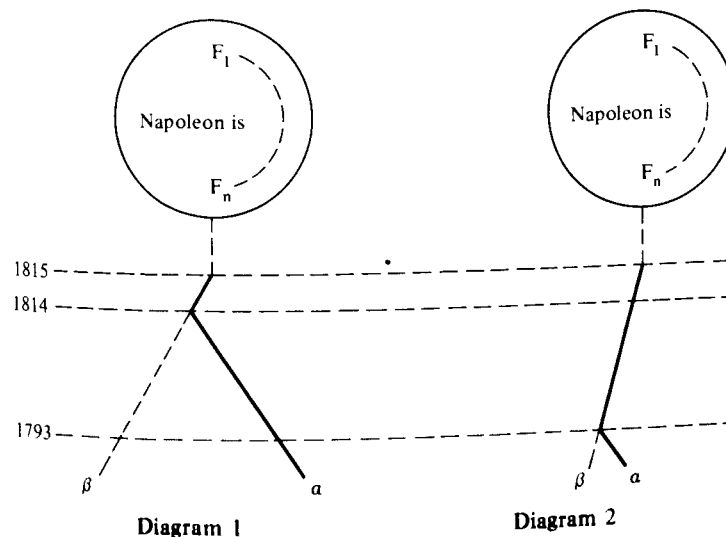
Suppose I get to know a man slightly. Suppose then a suitably primed identical twin takes over his position, and I get to know him fairly well, not noticing the switch. Immediately after the switch my dossier will still be dominantly of the original man, and I falsely believe, as I would acknowledge if it was pointed out, that *he* is in the room. Then I would pass through a period in which neither was dominant; I had not misidentified one as the other, an asymmetrical relation, but rather confused them. Finally the twin could take over the dominant position; I would not have false beliefs about who is in the room, but false beliefs about, for example, when I first met the man in the room. These differences seem to reside entirely in the differences in the believer's reactions to the various discoveries, and dominance is meant to capture those differences.

Dominance is not simply a function of *amount* of information (if that is even intelligible). In the case of persons, for example, each man's life presents a skeleton and the dominant source may be the man who contributed to covering most of it rather than the man who contributed most of the covering. Detail in a particular area can be outweighed by spread. Also the believer's reasons for being interested in the item at all will weigh.

Consider another example. If it turns out that an impersonator had taken over Napoleon's role from 1814 onwards (post-Elba), the cluster of the typical historian would still be dominantly of the man responsible for the earlier exploits

<sup>10</sup> The term is D. Kaplan's, see 'Quantifying In', in D. Davidson and J. Hintikka (eds), *Words and Objections* (Dordrecht: Reidel, 1969): I think there are clear similarities between my notion of a dominant source and notions he is there sketching. However I want nothing to do with vividness. I borrow the term 'dossier' from H. P. Grice's paper 'Vacuous Names' in the same volume.

( $\alpha$  in Diagram 1) and we would say that they had false beliefs about who fought at Waterloo. If however the switch had occurred much earlier, it being an unknown Army Officer being impersonated, then their information would be dominantly of the later man ( $\beta$  in Diagram 2). They did not have false beliefs about who was the general at Waterloo, but rather false beliefs about that general's early career.



I think we can say that *in general* a speaker intends to refer to the item that is the dominant source of his associated body of information. It is important to see that this will not change from occasion to occasion depending upon subject-matter. Some have proposed<sup>11</sup> that if in case 1 the historian says 'Napoleon fought skilfully at Waterloo' it is the imposter  $\beta$  who is the intended referent, while if he had said in the next breath '... unlike his performance in the Senate' it would be  $\alpha$ . This seems a mistake; not only was what the man said false, what he intended to say was false too, as he would be the first to agree; it wasn't Napoleon who fought skilfully at Waterloo.

<sup>11</sup> K. S. Donnellan, 'Proper Names and Identifying Descriptions', in Davidson and Harman (eds), *op. cit.*, p. 371.

With this background, then, we may offer the following tentative definition:

- 'NN' is a name of  $x$  if there is a community  $C$
1. in which it is common knowledge that members of  $C$  have in their repertoire the procedure of using 'NN' to refer to  $x$  (with the intention of referring to  $x$ )
  2. the success in reference in any particular case being intended to rely on common knowledge between speaker and hearer that 'NN' has been used to refer to  $x$  by members of  $C$  and not upon common knowledge of the satisfaction by  $x$  of some predicate embedded in 'NN'.<sup>12</sup>

(In order to keep the definition simple no attempt is made to cover the sense in which an unused institutionally approved name is a name.)

This distinction (between use-because-(we know)-we-use-it and use upon other bases) is just what is needed to distinguish dead from live metaphors; it seems to me the only basis on which to distinguish referential functioning of names, which may grammatically be descriptions, from that of descriptions.<sup>13</sup>

The definition does not have the consequence that the description 'the man we call "NN"' is a name, for *its* success as a referential device does not rely upon common knowledge that *it* is or has been used to refer to  $x$ .

Intentions alone don't bring it about that a name gets a denotation; without the intentions being manifest there cannot be the common knowledge required for the practice.

Our conditions are more stringent than Kripke's since for him an expression becomes a name just so long as someone has dubbed something with it and thereby caused it to be in common usage. This seems little short of magical. Suppose one of a group of villagers dubbed a little girl on holiday in

<sup>12</sup> For the notion of 'common knowledge', see D. K. Lewis, *Convention* (Cambridge, Mass.: Harvard University Press, 1969) and the slightly different notion in S. R. Schiffer, *Meaning* (Oxford: Clarendon Press, 1972). For the notion of 'a procedure in the repertoire' see H. P. Grice 'Utterer's Meaning, Sentence Meaning, Word Meaning', *Foundations of Language* (1968). Clearly the whole enterprise owes much to Grice but no commitment is here made to any specific version of the theory of communication.

<sup>13</sup> And if Schiffer is right much more as well—see *Meaning*, chap. V.

the vicinity 'Goldilocks' and the name caught on. However suppose that there were two identical twins the villagers totally fail to distinguish. I should deny that 'Goldilocks' is the name of either—even if by some miracle each villager used the name consistently but in no sense did they fall into two coherent sub-communities. (The name might denote the girl first dubbed if for some peculiar reason the villagers were deferential to the introducer of the name—of this more below.)

Consider the following case. An urn is discovered in the Dead Sea containing documents on which are found fascinating mathematical proofs. Inscribed at the bottom is the name 'Ibn Khan' which is quite naturally taken to be the name of the constructor of the proofs. Consequently it passes into common usage amongst mathematicians concerned with that branch of mathematics. 'Khan conjectured here that . . .' and the like. However suppose the name was the name of the scribe who had transcribed the proofs much later; a small '*id scripsit*' had been obliterated.

Here is a perfect case where there is a coherent community using the name with the mathematician as the intended referent and a consequence of the definition would be that 'Ibn Khan' would be one of his names. Also, 'Malachi' would have been the name of the author of the biblical work of the same name despite that its use was based upon a misapprehension ('Malachi' means my messenger).<sup>14</sup>

Speakers within such traditions use names under the misapprehension that their use is in conformity with the use of other speakers referring to the relevant item. The names would probably be withdrawn when that misapprehension is revealed, or start a rather different life as 'our' names for the items (cf. 'Deutero Isaiah' etc.). One might be impressed by this, and regard it as a reason for denying that those within these traditions spoke the literal truth in using the names. It is very easy to add a codicil to the definition which would have this effect.

Actually it is not a very good reason for denying that speakers within such traditions are speaking the literal

<sup>14</sup> See O. Eissfeldt, *The Old Testament: An Introduction* (Oxford: Oxford University Press, 1965), p. 441.

truth.<sup>15</sup> But I do not want to insist upon any decision on this point. This is because one can be concessive and allow the definition to be amended without giving up anything of importance. First: the definition with its codicil will still allow many names to change their denotation. Secondly: from the fact that, in our example, the community of mathematicians were not denoting the mathematician it obviously fails to follow that they were denoting the scribe and were engaged in strictly speaking massive falsehood of him.

Let me elaborate the first of these points.

There is a fairly standard way in which people get their names. If we use a name of a man we expect that it originated in the standard manner and this expectation may condition our use of it. But consider names for people which are obviously nicknames, or names for places or pieces of music. Since there is no standard way in which these names are bestowed subsequent users will not in general use the name under any view as to its origin, and therefore when there is a divergence between the item involved in the name's origin and the speakers' intended referent there will be no *mis*-apprehension, no latent motive for withdrawing the name, and thus no bar to the name's acquiring a new denotation even by the amended definition. So long as they have no reason to believe that the name has dragged any information with it, speakers will treat the revelation that the name had once been used to refer to something different with the same sort of indifference as that with which they greet the information that 'meat' once meant groceries in general.

We can easily tell the story in case 2 of our Napoleon diagram so that  $\alpha$  was the original bearer of the name 'Napoleon' and it was transferred to the counterfeit because of the similarity of their appearances and therefore without the intention on anyone's part to initiate a new practice. Though this is not such a clear case I should probably say that historians have used the name 'Napoleon' to refer to  $\beta$ . They might perhaps abandon it, but that of course fails to show that they were all denoting  $\alpha$ . Nor does the fact that someone in the know might come along and say 'Napoleon

<sup>15</sup> John McDowell has persuaded me of this, as of much else. He detests my conclusions.

was a fish salesman and was never at Waterloo' show anything. The relevant question is: 'Does this contradict the assertion that was made when the historians said "Napoleon was at Waterloo"?' To give an affirmative answer to this question requires the prior determination that they have all along been denoting  $\alpha$ .

We need one further and major complication. Although standardly we use expressions with the intention of conforming to the general use made of them by the community, sometimes we use them with the *overriding* intention to conform to the use made of them by some other person or persons. In that case I shall say that we use the expression *deferentially* (with respect to that other person or group of persons). This is true of some general terms too: 'viol', 'minuet' would be examples.

I should say, for example, that the man in the conversation in the pub used 'Louis' deferentially. This is not just a matter of his ignorance; he could, indeed, have an opinion as to who this Louis is (the man he met earlier perhaps) but still use the expression deferentially. There is an important gap between

intending to refer to the  $\phi$  and believing that  $a =$  the  $\phi$ ;

intending to refer to  $a$

for even when he has an opinion as to who they are talking about I should say that it was the man they were talking about, and not the man he met earlier, that he intended to refer to.

Archaeologists might find a tomb in the desert and claim falsely that it is the burial place of some little known character in the Bible. They could discover a great deal about the man in the tomb so that he and not the character in the Bible was the dominant source of their information. But, given the nature and point of their enterprise, the archaeologists are using the name deferentially to the authors of the Bible. I should say, then, that they denote that man, and say false things about him. Notice that in such a case there is some point to this characterization.

The case is in fact no different from any situation in which a name is used with the overriding intention of referring to

something satisfying such-and-such a description. Kripke gives the example of 'Jack the Ripper'. Again, after the arrest of a man *a* not in fact responsible for the crimes, *a* can be the dominant source of speakers' information but the intended referent could well be the murderer and not *a*. Again this will be productive of a whole lot of falsehood.

We do not use all names deferentially, least of all deferentially to the person from whom we picked them up. For example, the mathematicians did not use the name 'Ibn Khan' with the *overriding* intention of referring to whoever bore that name or was referred to by some other person or community.

We must thus be careful to distinguish two reasons for something that would count as 'withdrawing sentences containing the name'

- (a) the item's not bearing the name 'NN' ('Ibn Khan', 'Malachi')
- (b) the item's not being NN (the biblical archaeologists).

I shall end with an example that enables me to draw these threads together and summarize where my position differs from the Causal Theory.

A youth *A* leaves a small village in the Scottish highlands to seek his fortune having acquired the nickname 'Turnip' (the reason for choosing a nickname is I hope clear). Fifty or so years later a man *B* comes to the village and lives as a hermit over the hill. The three or four villagers surviving from the time of the youth's departure believe falsely that this is the long-departed villager returned. Consequently they use the name 'Turnip' among themselves and it gets into wider circulation among the younger villagers who have no idea how it originated. I am assuming that the older villagers, if the facts were pointed out, would say 'It isn't Turnip after all' rather than 'It appears after all that Turnip did not come from this village.' In that case I should say that they use the name to refer to *A*, and in fact, denoting him, say false things about him (even by uttering 'Here is Turnip coming to get his coffee again').

But they may die off, leaving a homogeneous community

using the name to refer to the man over the hill. I should say the way is clear to its becoming his name. The story is not much affected if the older villagers pass on some information whose source is *A* by saying such things as 'Turnip was quite a one for the girls', for the younger villagers' clusters would still be dominantly of the man over the hill. But it is an important feature of my account that the information that the older villagers gave the younger villagers could be so rich, coherent, and important to them that *A* could be the dominant source of their information, so that they too would acknowledge 'That man over the hill isn't Turnip after all.'

A final possibility would be if they used the name deferentially towards the older villagers, for some reason, with the consequence that no matter who was dominant they denote whoever the elders denote.

6. *Conclusion.* Espousers of both theories could reasonably claim to be vindicated by the position we have arrived at. We have secured for the Description Theorist much that he wanted. We have seen that for at least the most fundamental case of the use of names (non-deferentially used names) the idea that their denotation is fixed in a more or less complicated way by the associated bodies of information that one could cull from the users of the name turns out not to be so wide of the mark. But of course that the fix is by causal origin and not by fit crucially affects the impact this idea has upon the statement of the truth conditions of existential or opaque sentences containing names. The theorist can also point to the idea of dominance as securing what he was trying, admittedly crudely, to secure with his talk of the 'majority of' the descriptions, and to the 'degree of fit requirement' as blocking consequences he found objectionable.

The Causal Theorist can also look with satisfaction upon the result, incorporating as it does his insight about the importance of causality into a central position. Further, the logical doctrines he was concerned to establish, for example the non-contingency of identity statements made with the use of names, are not controverted. Information is individuated by source; if *a* is the source of a body of information

nothing else could have been. Consequently nothing else could have been *that a*.

The only theorists who gain no comfort are those who, ignoring Kripke's explicit remarks to the contrary,<sup>16</sup> supposed that the Causal Theory could provide them with a totally *non-intentional* answer to the problem posed by names. But I am not distressed by their distress.

Our ideas also point forward; for it seems that they, or some close relative, must be used in explaining the functioning of at least some demonstratives. Such an expression as 'That mountaineer' in 'That mountaineer is coming to town tonight' may avert to a body of information presumed in common possession, perhaps through the newspapers, which fixes its denotation. No one can be *that* mountaineer unless he is the source of that information no matter how perfectly he fits it, and of course someone can be that mountaineer and fail to fit quite a bit of it. It is in such generality that defence of our ideas must lie.

But with these hints I must leave the subject.

<sup>16</sup> Kripke, *op. cit.*, p. 302.

## Identity and Predication

A translation is one thing, a theory of meaning another. A manual of translation aims to provide, for each sentence of the language under study, a way of arriving at a quoted sentence of another language which has the same meaning. A theory of meaning, on the other hand, entails, for each sentence of the language under study, a statement of what it means. A translator states no semantical truths at all, nor has he any need of the concepts of truth, denotation, and satisfaction. Semantical truths relate expressions to the world, and can be stated only by using, not mentioning, expressions of some language or other. Such truths arise inevitably in the construction of a theory of meaning, for statements of what a sentence means can be made only by using an equivalent sentence of the theory's language, and, if the theory is to be finite, must be deduced from statements assigning semantical properties to its parts.

The two activities proceed under quite different constraints. Certainly a manual of translation must be finite, deriving translations for sentences from correlations of the sentences' parts, but a translator may use any segmentation of sentences he finds convenient. The finiteness demand upon a theory of meaning, on the other hand, goes to the very heart of what it is a theory of, since a language for which there is no finite theory of meaning is an unlearnable language.<sup>1</sup> And the semanticist aims to uncover a structure

From *The Journal of Philosophy* 72, 13 (17 July 1975), pp. 343-63. Reprinted by permission. This is a revised version of a paper read at Princeton University and to the 1974 York Conference on Logic and Semantics. I would like to thank H. Margain, S. Schiffer, B. Stroud, and B. Taylor for their help, and to acknowledge an especial debt to conversation with J. H. McDowell and to the writings of D. Davidson.

<sup>1</sup> See D. Davidson, 'Theories of Meaning and Learnable Languages', in Y. Bar-Hillel (ed.), *Logic, Methodology, and Philosophy of Science* (Amsterdam: North-Holland, 1965). These points on the relation between translation and the theory of meaning are due to Davidson.