P.N. JOHNSON-LAIRD AND A. GARNHAM

DESCRIPTIONS AND DISCOURSE MODELS

I

Since the work of Frege and Russell, the literature on descriptions has been largely concerned with an analysis of the relations between words and the world. In this paper, we present an account based on a different approach with a mentalistic emphasis: we intend to show how knowledge and intentions underlie the use and interpretation of descriptions. In pursuing this analysis, we have been forced to postulate a crucial role for what we call *mental models of discourse*, or more simply, *discourse models*. A discourse model is a mental object that constitutes an individual's knowledge of a discourse. It is constructed on the basis of what has occurred in the discourse supplemented by general and specific knowledge. Questions may be answered by recourse to the information that it contains; assertions may be evaluated by reference to it; utterances and actions may be based upon it. We do not suppose that mental models are constructed only from discourse: they are also created from memory, perception, imagination, and the operation of other mental processes. Our concern, however, is with those models derived from discourse and with elucidating their relations to descriptions. Our starting point has a distinct psychological flavor, the well known contrast drawn by Keith Donnellan (1966) between referential and attributive uses of definite descriptions.

When a speaker uses a definite description *referentially*, Donnellan claimed, his intention is to enable a listener to pick out whom or what he is talking about. When a speaker uses a definite description *attributively*, he may have no specific individual in mind: his intention is to designate whoever or whatever satisfies his description. Donnellan emphasized that the distinction depends, not on syntax or semantics, but on a speaker's intentions: it is impossible to determine the function of a definite description by considering a sentence in isolation from its context. He also claimed that it was necessary to use a description in order to make an attributive designation. Hence, when a speaker uses a description attributively, the relevant entity must be accurately described. If someone claims:

 (1)   The murderer of Smith is insane

but Smith died of natural causes, then his speech act is abortive. But,

when a description is used referentially, there is no necessary demand
that the entity designated should fit the description. If a speaker asserts:

(2)   The woman drinking the martini is a publisher

then he may designate a particular individual, and be correctly under-
stood, even if that person is actually drinking vichy water and accord-
ingly fails to fit the description.

Donnellan's distinction in usage was overlooked by Frege, Russell,
Strawson, and other earlier workers, whose theories are consequently
deficient. It constitutes a definite step in the direction of a more
mentalistic analysis, but in our view it does not go far enough. We will
argue that Donnellan has not got the referential-attributive distinction
quite right: he has compared extraordinary referential uses of descrip-
tions with ordinary attributive uses. In the next section, we produce a
counter-example to Donnellan's claim that an attribution can only work
if the description is accurate. Its analysis leads us to introduce the
notion of a discourse model, which we use in Section IV to solve the
problem of the predicative uses of definite and indefinite descriptions.
Section V re-examines the referential-attributive distinction. Donnellan's
analysis is based on cases where if the speaker is able to pick out an
entity on the basis of a description, then so, too, is the listener. This
symmetry is not invariable, and in this section we make a systematic
exploration of cases in which it does not hold. They lead us to an
alternative formulation in Section VI of the referential-attributive dis-
tinction, which we treat as a dichotomy rather than a continuum. Finally,
we attempt to draw some general conclusions.

II

The flaw in Donnellan's account of attributive usage can best be
established by way of an example.[1] Let us suppose that a friend of yours
has been to the cinema and you are talking to his wife about the film he
saw. You might ask her the following question:

(3)   How did John enjoy the film he saw last night?

even if you do not know which particular film he saw. Likewise, John's
wife might reply truthfully:

(4)   He enjoyed it.

even if she does not know which particular film he saw, either. Both of you are plainly utilizing attributive designations. But, now consider a rather different case. John has told everyone that he is going to see *The Sound of Music*. He does indeed see it, but later he makes a clandestine visit to another cinema. You believe that unbeknownst to him his wife has found out about this visit to a second film. Even though neither she nor you know what the second film was, you might ask her:

(5)   How do you think John enjoyed the film that he doesn't know that you know he saw last night?

Your attributive designation will be perfectly intelligible to John's wife, for whom it is also attributive.

Let us now suppose that John's wife has discovered that her husband has found out about this visit to a second film. Even though neither she *striking fact that your question will still be perfectly intelligible to her.* Donnellan stresses that a speaker using a definite description attributively fails to say anything if nothing fits the description. But this example is a clear case to the contrary: the attributive description does not correctly describe the entity it designates. Nevertheless, the attribution works because John's wife is able to reconstruct what you believe to be the case, namely, that there is a film that her husband does not know that she knows he saw. She knows that in reality there is no such film; but she also knows that your supposition is entirely compatible with a world in which her husband is ignorant of her knowledge.

Other theories that incorporate Donnellan's account appear to be similarly vitiated by this counter-example (see, e.g. Stalnaker, 1972; Kaplan, 1977; Wilson, 1978). Kripke (1977), for example, argues – correctly, in our view – that the distinction between a referential and an attributive description is pragmatic rather than semantic. He claims that the 'semantic referent' of a non-indexical designator is given by the speaker's *general* intentions, whereas the 'speaker's referent' is given by his *specific* intention on a given occasion to refer to a certain object. In a referential use of a designatory expression, the speaker's referent will be the same as the semantic referent *if* the speaker's beliefs are correct, but they may not be the same if he is mistaken. In an attributive use of a designatory expression, the speaker's specific intention *is* simply to refer to the semantic referent. In this case, Kripke writes (p. 264): "the speaker's referent is, *by definition*, the semantic referent". The class of counter-examples that we have illustrated shows that this analysis of an attributive description is false.

III

Philosophers and linguists often talk of *the* context of an utterance. But such talk can be dangerously misleading: it is too easy to lose sight of another point that our counter-example to Donnellan sustains. There is usually one context for the speaker and another context for the listener. Indeed, it is difficult to account for the consummation of designatory acts unless one is prepared to grant that the real context of an utterance consists of separate representations of the current conversation that the speaker and the listener create and maintain. But what form do such representations take? This question is clearly an empirical one, but we can at least begin to frame an answer.

An initially plausible conjecture is that discourse is represented by expressions in some 'language' of the mind. But, although such representations may well exist, there are strong grounds for supposing that they must at the very least be supplemented by another sort of representation that provides an interpretation of expressions in the mental language. One line of argument comes from the way in which people ordinarily reason with quantified assertions (see Johnson-Laird, 1975; Johnson-Laird and Steedman, 1978). They do not possess a mental logic with rules of inference that apply to expressions in a language of thought. On the contrary, the evidence suggests that they know only the fundamental semantic principle governing logical necessity: an inference is valid if and only if there is no interpretation of the premises consistent with a denial of the conclusion. This knowledge gives rise to an inferential procedure that we will illustrate by an example.

Suppose you are told that of a group of individuals in a room:

(6)   All the authors are bachelors

then according to the present theory you represent this information by constructing a mental model in which there is some arbitrary number of mental tokens designating the authors and, in accordance with this premise, each of them is identified by another token as being a bachelor. Since there may be bachelors in the room who are not authors, you add further tokens to your model in order to represent them. You have accordingly created a mental model of the following sort:

(7)   author = bachelor
       author = bachelor
       author = bachelor
             (bachelor)
             (bachelor)
             (bachelor)

where the parentheses indicate that the relevant individuals may, or may not, exist. In order to *make* any inference, it is always necessary to go beyond logic in order to single out one conclusion from the potentially infinite number of valid, though generally trivial, conclusions that can be drawn from any set of premises. The evidence suggests that a common heuristic is to seek to maximize the number of identities on the fewest number of individuals in a model. Hence, if you are subsequently told:

(8)   All the churchgoers are bachelors

then the inferential heuristic will lead you to construct the following model:

(9)   author = bachelor = churchgoer
        author = bachelor = churchgoer
        author = bachelor = churchgoer
             (bachelor)
             (bachelor)
             (bachelor)

At this point, if you are lacking in logical prudence, you might be tempted to conclude that *all the authors are churchgoers* (or its converse). A small proportion of our subjects do, in fact, draw this fallacious conclusion. The majority of people, however, though they do not possess a mental logic, do grasp the fundamental semantic principle governing logical validity. They accordingly attempt to destroy the model giving rise to the putative conclusion without doing violence to the meaning of the premises. If you adopt this procedure, then you will readily discover that you can modify one of the identities without violating the premises:

(10)   author = bachelor = churchgoer
         author = bachelor = churchgoer
         author = bachelor
                  bachelor = churchgoer
              (bachelor)
              (bachelor)

You might now be tempted to conclude that *some of the authors are churchgoers* (or its converse); again, a proportion of subjects do draw such conclusions. But, if you persevere with the test, you will find that all the identities between authors and churchgoers can be destroyed:

(11)   author = bachelor
         author = bachelor
         author = bachelor
                 bachelor = churchgoer
                 bachelor = churchgoer
                 bachelor = churchgoer
                 (bachelor)

It should now be clear, as it is to about half the university students that we have tested, that there is no valid categorical conclusion to be drawn relating authors and churchgoers.

There is nothing mysterious about reasoning by the manipulation of models: the theory has been implemented in the form of a computer program. There are also strong grounds for supposing that people do construct such models, which represent individuals, events, properties, and relations. First, a perceptual representation of the world contains such entities, though in models that are very much richer in detail. Second, such a representation is powerful enough for an interpretation of multiply-quantified assertions. Third, the differences between such quantifiers as *each* and *every*, and *any* and *all*, can be readily accounted for, as Janet Fodor (1979) has independently shown, by recourse to such models. Fourth, several aspects of the psychological semantics of natural language are naturally elucidated by reference to models, e.g. the mental representation of truth conditions (see Johnson-Laird, 1979).

A theory of mental models must embody both a system of representing information and a set of procedures for constructing, manipulating, and interrogating such representations. An important aspect of the present theory is that individuals can be represented as unique even when no individuating information is known about them. Thus, the representation above of 'all the authors are bachelors' treats each token as designating a separate, potentially distinguishable, individual. If it is subsequently learnt that one of the authors has red hair, then this information can be represented in the model by selecting a token designating an author and linking it with one designating a red-haired individual. The link, of course, will represent an identity: the two individuals are to be taken as one and the same. Those people with recourse to vivid imagery may actually imagine a red-haired author, but the phenomenal experience is not crucial since many human beings appear to be bereft of images. The representation of individuals already *known* to be unique is a simple extension of the general principles. The red-haired author is, of course, unique in the discourse model, but

nothing is presumably known to distinguish him from the set of red-haired authors at large. On the one hand, a listener may have sufficient specific knowledge to individuate someone to whom reference is made, and can accordingly introduce a token into the discourse model that would be individuating in any such model by virtue of the particular pattern of links to other tokens, etc. On the other hand, a listener may know merely that a reference is *intended* to be individuating, but be able only to introduce a token that is individuating within the current discourse model.

Discourse models represent not merely the relevant individuals, events, relations, and so on, but also what is known about the knowledge of other participants. A speaker structures his remarks partly on the basis of what he knows about the listener's discourse model; a listener interprets utterances partly on the basis of what he knows about the speaker's discourse model. For instance, theorists often take for granted that a definite description denotes a unique individual, and pay little more than lip-service to the fact that descriptions which occur in actual conversations, if considered out of context, are seldom applicable to one and only one individual. They argue that context will take care of such problems. In fact, this assumption, though true, is an oversimplification. As Isard (1974) has emphasized in a paper that has considerably influenced our thoughts about context, communications do not merely depend on the context for their interpretation, they change that context. Linguists and philosophers have often noted that a definite description can lead a listener to infer the existence of a unique entity if the description occurs in the absence of a prior identification of the entity. They have seldom noted that such a description often establishes uniqueness only with respect to the current discourse model.[2] If a speaker remarks:

(12)   The man who lives next door to me has bought a birdbath

then it would be a foolhardy listener indeed who took this referential usage to entail (*pace* Russell, 1905), or to presuppose (*pace* Strawson, 1950), that there is one and only one man living next door to the speaker. No such claim is being made. The definite description in (12) designates the only neighbour of the speaker who is (or who is going to be) relevant in the current context.

Uniqueness in a model rather than in reality is what controls the use and interpretation of definite descriptions. If a speaker is to communicate felicitously, then he must consider whether an entity will be unique in his listener's model. Utterances need seldom be more than

clues about how to change a discourse model: they depend for their interpretation on what a listener knows, but that interpretation in turn modifies or extends the discourse model. A discourse model is in part a surrogate for reality. Indeed, it is sometimes convenient to speak as if language were used to talk about discourse models rather than the world.

<div align="center">IV</div>

The notion of a discourse model allows us to elucidate the problem of the *predicative* uses of descriptions. A sentence such as:

(13)   Hugh Hambling is a teacher at Akenfield

makes use of an indefinite description that does not identify an individual (see Vendler, 1967) as is evident from the fact that it would be odd to continue:

(14)   The teacher at Akenfield comes from Norfolk.

The indefinite description is used *predicatively* rather than to designate an individual. Unfortunately, the existence of predicative uses of indefinite descriptions appears to play havoc with the semantics of the indefinite article. Wilson (1978), for example, claims that predicative uses are incompatible with the thesis that 'a' and 'an' have the logical force of quantifiers. He also argues that there are predicative uses of definite descriptions that fail to designate an individual, e.g.:

(15)   William Russ is the gravedigger.

How is the distinction between predicative and designative uses of descriptions to be explained and the alleged problems about giving an account of their semantics to be resolved? One view, proposed by Wilson (1978), is that descriptions, both definite and indefinite, have an underlying logical form in which they are always a proper part of predicates and always prefixed by some sort of copula verb. The predicate in turn requires a genuine singular term or some other appropriate expression as its argument. It follows that identifying uses of indefinites such as:

(16)   I shot a tiger

have an underlying logical form equivalent to:

(17)   I shot something which is a tiger

and, likewise, that referential uses of definite descriptions such as:

(18)   I shot the tiger

have an underlying logical form equivalent to:

(19)   I shot this thing which is uniquely a tiger.

And attributive uses of definite descriptions such as:

(20)   I want to shoot the fiercest tiger in India

have an underlying logical form equivalent to:

(21)   I want to shoot anything which is the fiercest tiger in India.

Linguists will appreciate that this solution bears a striking resemblance to Bach's (1968) proposal of deriving nouns from restrivtive relative clauses. Bach, too, points out how neatly such a scheme accommodates predicative uses of descriptions, which he terms 'predicate nominals'. Yet the approach is not without its difficulties. By using a universal quantifier in the analysis of attributives, there is no remaining machinery for specifying the underlying logic of a generic assertion based on a definite noun phrase, such as:

(22)   The tiger is a fierce animal.

Moreover, there is no very ready account of the relation between such sentences as:

(23)   Ernie is the driver

and

(24)   The driver is Ernie.

Are we to suppose that in (24) there is a predicative use of a proper name since, according to Wilson, the definite description cannot be a singular term? We prefer an alternative approach, based on discourse models, which yields a uniform account of designative and predicative descriptions, whether definite or indefinite.

The representation of a sentence containing a designative indefinite description:

(25)   Anne met an odd-job man

requires a token to be introduced into the discourse model to represent a

member of the class of odd-job men. The individuating token representing Anne is then linked to that token in order to represent the appropriate relation. The representation of a sentence containing a predicative indefinite description:

(26)   Hugh is a teacher at Akenfield

requires a relation of identity to be established between the individuating token representing Hugh and one of the tokens representing members of the class of teachers at Akenfield. Which of the two items, if any, has to be added *de novo* to the discourse model depends on what has already been established there. The assertion may lead to the listener assuming for the first time that there *are* teachers at Akenfield: a token is introduced to represent a member of that class, and it is linked to the token representing Hugh. Alternatively, the token representing Hugh may have to be added to the discourse model. It should be noted that Vendler's claim that predicative indefinite descriptions do not establish existence is slightly misleading. They may establish the existence of entities that match their descriptive content (e.g. teachers at Akenfield), depending on what the listener already knows, but this point is independent of the existence of the entity to which the predicate applies (e.g. Hugh).

The distinction between predicative and designative uses of definite descriptions can be handled in an analogous fashion. The designative description in the sentence:

(27)   Ralph talked to the gravedigger

requires a representation in the form of a token representing a member of the class of gravediggers that is either individuated by previous identifying information or is established by (27) as the unique member of the class of gravediggers relevant to the present context. The predicative use of a definite description:

(28)   William is the gravedigger

is represented by an identity between two unique tokens, one designating William and the other designating the gravedigger. Once more, which of these tokens, if any, has to be added *de novo* to the discourse model depends on what has already been established there.

In general, an indefinite description calls for one token of a class corresponding to its descriptive content to be linked to the other arguments, if any, of the verb, and for other such tokens not so linked to be specified as optional. The nature of the links between tokens depends

on the meaning of the verb. This analysis applies to both designative and predicative uses; and it is wholly independent of whether or not the existence of the entities represented has been previously established. The token that is picked out by a designative use normally identifies a new candidate for reference. The token that is picked out by a predicative use does not identify a new candidate for reference: the singular term to which the predicate is applied pre-empts that possibility. In either case, it is evident that contrary to a thesis that has been urged by a number of theorists (e.g. Stenning, 1977, 1978; Hawkins, 1978), an indefinite description does not necessarily have the force of an existential quantifier. In particular, the existence of the relevant entity may have been established independently:

(29)   There are some dangerous animals, including lions, in the zoo. Yesterday, one of the zoo's keepers was attacked by a lion there. The lion had to be destroyed.

Here, the indefinite description, 'a lion', in the second sentence merely singles out one of the lions whose existence is established by the first sentence. However, a listener may infer the existence of an entity corresponding to an indefinite description, even if it has not been previously established.

A singular definite description calls for one unique token corresponding to its descriptive content to be appropriately linked to other arguments of the verb, and it specifically debars the presence of other tokens of the same type from the discourse model. A plural definite description calls for each token of a unique set to be appropriately linked to other arguments of the verb. However, a plural description is often susceptible to two different interpretations. Thus, the sentence:

(30)   The men lifted the table

may be interpreted to mean either that each man lifted the table separately or that all the men lifted the table together. Some verbs ordinarily allow only one of the two possible interpretations:

(31)   The men sat round the table.

        The men stood on one leg and sang the 'Horst Wessel' song.

Hence, the link is either from the set as a whole or each of its members to the other arguments of the verb.

The predicative uses of descriptions are easily elucidated with the aid of discourse models; since they present no special difficulties, we will

concentrate in the remainder of this paper on designative uses, and on the attributive-referential distinction.

<center>V</center>

In our view, what ultimately underlies usage is a speaker's intentions. Speakers intend certain knowledge to be relevant to the interpretation of their utterances. It follows that the limitations of their knowledge place constraints on their intentions, and it can sometimes seem that it is knowledge *per se* that determines usage. This is an *illusion*. Intentions are the decisive lever, and this fact should be borne in mind in considering the difference between the referential and attributive uses of definite descriptions.

When Donnellan introduced the distinction, he failed to explore systematically the consequences of speakers and hearers having different discourse models, that is, having different knowledge or beliefs about the universe of discourse. Our aim now is to make such an exploration in order to examine the interplay between knowledge and intention. There are four basic cases to be considered, depending on whether a speaker intends to designate an entity referentially or attributively, and on whether he intends the listener to interpret his designation referentially or attributively.

(1) Let us suppose that you know that Arthur was given some flowers by a visitor, but you know nothing about who the visitor was. If you believe that your listener knows nothing of this event, but wishes to know where the flowers came from, then you might remark:

(32)   The flowers were given to Arthur by a visitor.

Your *indefinite description*, 'a visitor', is clearly used in an *identifying* manner, but neither you nor your listener has a particular person in mind. Both of you would be forced to designate the relevant individual by a description:

(33)   The visitor is plainly very rich because the flowers are rare orchids.

The definite description, 'the visitor', is here used in an attributive way.

(2) Let us now consider a rather different example, which we have borrowed from Wilson (1978). Tom, Dick and Harry are all Presidential Advisers, and they are discussing whether or not it will be safe to hide some presidential tapes in a particular room of the White House. Tom

and Dick alone have been told that the head of the Russian Secret Police will be making a clandestine visit to see the President, and that he will be staying in that room. Tom suspects that Dick has forgotten this fact, and he needs to alert him to it without arousing Harry's suspicions since he is not privy to the secret. Hence, Tom remarks:

(34)   A guest of the President will be staying in the room.

His intention as far as Dick is concerned is to use an indefinite description that refers to a specific individual. If his intentions are consummated, then Dick will indeed take the remark to refer to that individual. Hence, if Dick replies:

(35)   Yes, the guest is due in a month's time.

then his definite description will be used referentially both for him and for Tom, though as far as Harry is concerned it can pass as an innocent attribution. Of course, it is unusual to use indefinite descriptions in a way that is referential for both speaker and hearer: it is more natural to use a definite description directly. That is why a rather elaborate scenario is necessary to motivate the example.

Examples (1) and (2) illustrate cases of mutual knowledge (always excepting Harry), and hence what is attributive for the speaker is attributive for the listener, and what is referential for the speaker is referential for the listener. Obviously, there can be critical discrepancies in their knowledge, and it is important to appreciate that they need neither be a cause nor a result of misunderstanding. A speaker may deliberately intend to make it clear that there is a discrepancy between his and his listener's knowledge. These cases are particularly interesting and are illustrated in the following two examples.

(3) If you have reason to believe that your listener met someone, but you know nothing about the circumstances of the individual concerned, then you might remark:

(36)   I understand you met a lady last night.

Your indefinite description identifies an individual whom you can designate only attributively,[3] e.g.:

(37)   Tell me more about the lady you met

but whom you suppose your listener will be able to designate referentially:

(38)   Yes, I met the only lady to have a British civil airline pilot's licence.

(4) It is easy to see that a speaker may have exactly the converse intentions, as when he remarks:

(39)   I took a girl out to dinner yesterday.

Here, he identifies an individual to whom he can refer but whom he supposes the listener can designate only attributively, e.g.:

(40)   Tell me more about the girl.

The speaker can choose to maintain the asymmetry in knowledge:

(41)   I don't want to tell you anything about the girl.

He may even make the point by way of a paradox:

(42)   I don't want to tell you anything about the best clog-dancer in Cleethorpes.

This remark does communicate some information (or misinformation) about the girl – indeed, as we shall argue later, it puts the listener in the position of being able to refer to her – but it is clearly intended to be tantalizing, to whet the listener's appetite, and to establish the speaker's intention not to be wholly forthcoming. Alternatively, once having identified the girl, the speaker may choose to characterize her in more than sufficient detail for the listener to be able to refer to her (subsequently).

The four cases that we have distinguished may also occur within opaque contexts. Where an indefinite description is intended to identify an individual, as might be the case with:

(43)   Joan wants to meet a philosopher

then there is no essential difference in the analysis. Here, there is a specific philosopher whom Joan wants to meet and what is at stake is whether anyone, including the speaker or hearer, is sufficiently knowledgeable to be able to intend to refer to that philosopher. It is worth remarking that Joan herself may not be able to do so. She may, for instance, know only that she wants to meet the philosopher who turned down the Nobel prize for literature. In this case, the speaker of sentence (43) may share her ignorance or alternatively he may be able to (and intend to) designate the individual referentially:

(44)   She wants to meet the man who refused the Nobel prize for
         literature – Jean-Paul Sartre, though she doesn't realize it.

If Joan has no specific philosopher in mind, then any accurate use of an
indefinite description to describe her wants will be non-identifying. The
state of her mind constrains a speaker's intentions. Hence, if a speaker
wishes to use a definite description, then it must be attributive for both
him and his listener:

(45)   The philosopher must be able to tell her the meaning of life.

Clearly, such attributive specifications may, or may not, be satisfied by
an actual individual. However, the same possibilities apply even if Joan
has a specific individual in mind: in talking about someone's pro-
positional attitudes, a speaker talks of the entities within the relevant
universe of beliefs, desires, wants, and so on. Consider, for example, the
following discourse (with apologies to Paul Jennings):

(46)   Joan wants to meet a famous philosopher – Pierre Marie
         Ventre, the founder of Resistentialism and the author of that
         celebrated maxim, "Les choses sont contre nous." Frankly, I
         don't believe the man exists.

The definite description, 'the man', in the final sentence can hardly be
interpreted as attributive; the speaker is referring to a specific individual
within the universe of Joan's wants. There is no contradiction between
referential usage and non-existence in reality. We are accordingly led to
the conclusion that the referential-attributive distinction is independent
of questions of opacity. It follows that attempts to reduce one
phenomenon to the other cannot work (cf. Bell, 1973; Cole, 1978).
   We have described a distinction in usage that applies to both definite
and indefinite descriptions, and which depends on a speaker's com-
municative intentions. We add that the same distinction occurs in the
use of proper names (as was pointed out by, for example, Stalnaker, 1972)
and in the use of quantified noun phrases (as was pointed out by, for
example, Johnson-Laird, (1970). We have also illustrated how the com-
munication of such intentions depends on the contents of the speaker's and
hearer's models of discourse. Our examples establish that the referential-
attributive distinction has to be made not once, but twice. A speaker may
intend to say something about a particular entity that he has in mind, or he
may intend to say something about whatever fits a description that he uses.

But it is a separate question whether he intends his hearer to interpret his remarks as referring to a specific entity or merely to whatever fits the description. A speaker can intend one thing for himself and another for his audience.

IV

Discourse models, and the conventions governing discourse, can account for many aspects of the use and interpretation of descriptions. Unfortunately, the basic distinction between referential and attributive usage remains somewhat vague, since if one asks whether the two categories are exhaustive or merely the opposite ends of a continuum, no ready answer is forthcoming.

Barbara Partee (1972) suggests that there may indeed be a continuum from attribution to reference, and she bases this view on Kaplan's concept of the vividness of a singular term or name. Kaplan (1969) distinguishes three different aspects of a name, which he uses in the broad sense of a singular term: its descriptive content, which relates it to the world and determines what, if anything, it denotes; its genetic character, which accounts for how a given user of the name acquired the belief that there is some individual that it is a name or description *of* (cf. Kripke, 1972, who offers a similar causal account of proper names); and its vividness, which depends on the conglomeration of mental images and descriptions that its user employs to bring the relevant individual to mind. Kaplan's aim is to formulate the conditions under which one is warranted in moving from an assertion such as:

(47)   Holmes believes that the murderer of Smith is insane

to one in which the description is 'exported' from within the opaque context:

(48)   Holmes believes of the murderer of Smith that he is insane.

His thesis is that a necessary condition for this step is that the singular term is sufficiently vivid for Holmes. In short, it can only occur when the definite description is truly referential. Initially, the description may be just a dull attribution as far as Holmes is concerned but, as Kaplan puts it: "At some point in his investigation, the slow accretion of evidence, all 'pointing in a certain direction' may just push Holmes' description over the appropriate vividness threshold so that we *would* say that there is now someone whom Holmes believes to be the murderer." Although

Partee endorses this view, she points out that it is by no means clear where one crosses the threshold from attribution to reference.

In our view, this approach to distinguishing between the two sorts of usage introduces a subtle confusion: it replaces the speaker's intentions as the central notion by something only indirectly relevant, his knowledge. As Donnellan emphasized, the distinction in usage is not just a matter of a speaker's beliefs or knowledge. A speaker may know that Jones, who is for him a very vivid individual, drove your car, but if in commenting on the way your car was driven, he remarks:

(49)   The man who drove your car is a maniac

then his intention may be entirely attributive. After all, if the statement is true, it remains true even if Jones was not the driver (see Donnellan, 1966, for comparable examples). Yet, although knowledge is not crucial to the distinction, it is relevant because it places constraints on possible intentions. If all you know about an individual is that he drove your car, then you cannot intend, though you may pretend, to designate him referentially. It is obvious that all that you need to know in order to make an attributive designation is the equivalent of a single individuating description (in context). Hence, the crux of the matter is what do you need to know in order to be able to designate an individual referentially? One alternative to Kaplan and Partee's answer is that you should have sufficient knowledge of the relevant individual to be able to designate that individual in more than one independent way. This principle plainly replaces the continuum of usage with a simple dichotomy. It has the advantage of simplicity, but is it correct?

Let us approach the problem from a slightly different tack. If you designate an individual referentially, then you should be able to pick out that individual for yourself. You should know who that individual is: you should be able to answer a question of the form *who is the so-and-so*? Kaplan's view of what it takes to answer such a question would be that you should possess a sufficiently 'vivid' description of the individual. This view, however, is not one that holds for ordinary conversations. If you are asked a question such as:

(50)   Who was the first man to see a living human retina?

then you might well reply:

(51)   The man who invented the ophthalmoscope.

On the one hand, your reply may rest on your knowledge both that there was an individual who invented the ophthalmoscope and that he was

thus able to see a living human retina for the first time. If you intended to refer to this individual, then your use of the definite description was referential. On the other hand, your answer to the question might simply reflect an ingenious Holmesian deduction. You argued to yourself that the only way to see a living retina is by using an ophthalmoscope, that the person who invented it was presumably the first person to see such a retina, that the instrument was probably invented in the nineteenth century, and hence the inventor was likely to be a man because there were few women scientists in those days. In this case, your description is clearly attributive rather than referential.

The points to be emphasized are that in order to use a description referentially you need to possess sufficient information to designate the relevant individual in more than one independent way, and that the independence is a matter not just of logic but also of epistemology. It hinges on how the speaker came by this knowledge: if he derived a second description from his original description by some process of thought, perhaps of a quasi-deductive nature, then his descriptions fail to be independent.

Two independent individuating descriptions seem to be a fairly minimal requirement, but is the criterion perhaps too much? Can a single description suffice for referential usage? Suppose that you possess the information that there once was a journalist who gave solo unaccompanied concerts in which he sang complete operettas, and that no other journalist ever did so. It is clear that you can use your knowledge to designate this individual attributively – no amount of knowledge is ever a bar to speaking attributively. What is less clear is whether you can intend to refer to this individual and, if so, how you could go about doing so. Obviously, in order to make an informative assertion about him, you need to know at least some other property that he possessed. You might, for example, be able to assert:

(52)   The journalist who sang solo operettas was Viennese.

This information is plainly not individuating, and it would not suffice to enable you to refer to the journalist. Hence, if you were asked:

(53)   Who was the journalist who sang solo operettas in concerts?

then an answer of the form:

(54)   The Viennese journalist

would not be adequate. The answer fails to provide an individuating designation even in context. Indeed, the use of a definite description

here is rather revealing: it suggests that your description is elliptical and that you are able to provide further genuinely individuating information. As a matter of fact, you are unable to do so.

A more pertinent minimal requirement for reference is a knowledge of a single individuating description together with the appropriate proper name. In ordinary discourse, if you are asked:

(55)   Who is the man who invented holography?

then you would be happy to reply:

(56)   Denis Gabor

even if you knew nothing else about him. Such an answer can hardly be considered to be attributive: you had a specific individual in mind and intended to refer to him. A name is not a description, but it is individuating. Yet, a name alone is not the passport to being able to refer to an individual, since names too can be used attributively. Such may be the intention of a speaker who asks, 'Who *is* Paul Jennings?' (see Stalnaker, 1972, for further arguments in support of the notion of an attributive use of proper names). These considerations lead us to conclude that a knowledge of two independent ways of individuating an entity is a cognitive prerequisite for an intention to refer to that entity, and that the referential-attributive distinction is a dichotomy rather than a continuum.

VII

We have argued that the intentions of a speaker, his knowledge (or beliefs), including his knowledge (or beliefs) about what his audience knows, are all of them important in determining the kind of use to which a description has been put. In this final section, we will try to give a more general account of these requirements. It will be necessary to bear in mind that the referential-attributive distinction can be drawn independently for speakers and hearers, so various combinations of the sets of beliefs outlined are possible. For the sake of simplicity, our account is solely in terms of assertions about the properties of entities, considered from the standpoint of their representation in discourse models.

If a speaker knows that there is one and only one member, f, of the class F such that f is a member of the class G, then the speaker may identify this individual using an assertion of the form:

(57) $\ulcorner$An F is $G$$\urcorner$

Later in the same discourse he may designate f using a definite description of the form:

(58) $\ulcorner$The F (who is G)$\urcorner$

where the parentheses indicate an optional restrictive relative clause (cf. Vendler, 1967).

When an indefinite description is used to identify an entity the speaker assumes that someone, not necessarily either the speaker or the hearer, can provide at least one unique description of that object. He intends that this person should be able to substitute the description(s) into statements about the object, *salva veritate*. That is to say, he intends his statement to be true only in possible worlds[4] in which the object designated has the same properties it has in the actual world.

In contrast, a non-identifying use of an indefinite description, for example:

(59) $\ulcorner$John wants an $F$$\urcorner$

does not commit the speaker to the assumption that someone is able to provide a unique description of any object which is an F. Other properties of Fs, including unique properties which a particular F might have, are irrelevant to whether (59) is true.

Predicative indefinites are a special case of identifying uses. A unique description fitting the object of which the predication is made is present in the utterance, and is, therefore, available to both speaker and hearer.

Referential and attributive uses of definites are distinguished by the knowledge which a speaker intends to be relevant to the interpretation of his utterance. When (58) is used attributively, no other unique descriptions which fit the designation are relevant to its interpretation, even if they are known to the speaker. He intends his utterance:

(60) $\ulcorner$The F (who is G) is $H$$\urcorner$

to be true in all possible worlds in which the entity designated by (58) has the property H whatever other properties it may have. In any given world this entity may or may not be the same one as is designated by the description in the actual world.

In order to use a description referentially, a speaker must know of its designation, f, that it is both the one and only F which is G, and either the one and only member of at least one other class J, or has the proper name $\ulcorner$A$\urcorner$. If he asserts (60), then he intends that:

(61) $\ulcorner A$ is $H\urcorner$

and/or statements of the form:

(62) $\ulcorner$ The J is $H\urcorner$

should be true. He is committed to substitution *salva veritate* of other designations of *f* into (60). That is to say, he takes (60) to be true only in worlds in which the entity designated by (58) has the properties and/or name which it has in the actual world, and in which it also has the property *H*.

In conclusion, let us go to the scene of a rather different crime from the one so vividly sketched in Donnellan's case-book. Brown is lying on the ground, apparently dead, his body strangely mutilated. The great detective examines his injuries and announces to the assembled onlookers:

(63)   Brown's murderer is a left-handed seamstress addicted to caraway seeds.

For once, Holmes is totally mistaken. Brown has not been murdered: his wounds are not mortal and he survives the attack. Yet, although Holmes's designation is technically erroneous, his speech act is hardly vacuous. The police will be looking for an assailant who fits the predication. A speaker such as Holmes intends that an attributive description should be satisfied by some entity, and that no other aspect of that entity is relevant to the designation. But these criteria are absolute only in intention, not reality. The speaker's assumptions may be mistaken and his intentions thwarted. It will make little difference to his designatory act provided that his listeners are able to construct a new attributive description on the basis of their knowledge, and provided that this description is satisfied by some entity. We are led ineluctably to the conclusion that whether or not a designation is consummated depends upon the speaker's and the listener's model of the discourse. A major intellectual task facing theorists is to establish a relation between such models and the models of formal semantics.

Centre for Research on Perception and Cognition,
Laboratory of Experimental Psychology,
University of Sussex.

NOTES

[1] We first became aware of such cases in trying to develop some counter-examples to a hypothesis about the role of mutual knowledge in the use of definite descriptions proposed by Clark and Marshall (1978).

[2] The notion of a discourse model has, of course, a number of different antecedents. It is clearly presaged by Karttunen's paper on discourse referents, which circulated some years ago (see Karttunen, 1976). Recently, Stenning (1977, 1978), Webber (1978), and others, have been led to postulate the existence of discourse models similar in outline to our own conception.

[3] We note in passing that such sentences as, 'Tell me something about a particular person that you have met', indicate that the occurrence of such expressions as 'a particular person' and 'a certain person' does not invariably signal that the speaker is able to refer to the relevant person (*pace* Partee, 1972; Wilson, 1978).

[4] We use the term 'possible world' for purely expository reasons. We do not necessarily wish to commit ourselves to the machinery of possible worlds semantics as developed, for example, by Kripke (1963).

REFERENCES

Bach, E.: 1968, 'Nouns and noun phrases'. In E. Bach and R.T. Harms (Eds.), *Universals in Linguistic Theory*, Holt, Rinehart, and Winston, New York.

Bell, J.: 1973, 'What is referential opacity?', *Journal of Philosophical Logic* 2, 155–180.

Clark, H.H. and Marshall, C.: (1978), 'Definite reference and mutual knowledge'. A paper presented at The Sloan Workshop on Computational Aspects of Linguistic Structure and Discourse Setting, University of Pennsylvania, May, 1978.

Cole, P.: 1978, 'On the origins of referential opacity.' In P. Cole (Ed.), *Syntax and Semantics, Vol. 9: Pragmatics*, Academic Press, New York.

Donnellan, K.: 1966, 'Reference and definite descriptions', *Philosophical Review* 75, 281–304.

Fodor, J.D.: 1979, 'The mental representation of quantifiers', Paper delivered at Symposium on Formal Semantics and Natural Language, University of Texas at Austin.

Hawkins, J.A.: 1978, *Definiteness and Indefiniteness: A Study in Reference and Grammaticality Prediction*, Croom Helm, London.

Isard, S.: 1974, 'Changing the context'. In E.L. Keenan (Ed.), *Formal Semantics of Natural Language*. Cambridge University Press, Cambridge.

Johnson-Laird, P.N.: 1970, 'The interpretation of quantified sentences'. In G.B. Flores d'Arcais and W.J.M. Levelt (Eds.), *Advances in Psycholinguistics*. North-Holland, Amsterdam.

Johnson-Laird, P.N.: (1975), 'Models of deduction'. In R.J. Falmagne (Ed.), *Reasoning: Representation and Process in Adults and Children*. Erlbaum, Hillsdale, N.J.

Johnson-Laird, P.N.: 1979, 'Formal semantics and the psychology of meaning', Paper delivered at Symposium on Formal Semantics and Natural Language, University of Texas at Austin.

Johnson-Laird, P.N. and Steedman, M.: 1978, 'The psychology of syllogisms', *Cognitive Psychology* **10**, 64–99.

Kaplan, D.: 1969, 'Quantifying in', In D. Davidson and K.J.J. Hintikka (Eds.), *Words and Objections: Essays on the Work of W.V.O. Quine*, Reidel, Dordrecht, Holland.

Kaplan, D.: 1977, 'Demonstratives: an essay on the semantics, logic, metaphysics and epistemology of demonstratives and other indexicals', Paper prepared for the Symposium on Demonstratives at a meeting of the Pacific Division of the American Philosophical Association, March, 1977.

Karttunen, L.: 1976, 'Discourse referents'. In J.D. McCawley (Ed.), *Syntax and Semantics 7: Notes from the Linguistic Underground*, Academic Press, New York.

Kripke, S.A.: 1963, 'Semantic considerations on modal logic', *Acta Philosophica Fennica* **16**, 83–94.

Kripke, S.A.: 1972, 'Naming and necessity'. In D. Davidson and G. Harman (Eds.), *Semantics of Natural Language*, Reidel, Dordrecht, Holland.

Kripke, S.A.: 1977, 'Speaker's reference and semantic reference', *Midwest Studies in Philosophy II*, 255–276.

Partee, B.H.: 1972, 'Opacity, coreference, and pronouns'. In D. Davidson and G. Harman (Eds.), *Semantics of Natural Language*, Reidel, Dordrecht, Holland.

Russell, B.A.W.: 1905, 'On denoting', *Mind* **14**, 479–493.

Stalnaker, R.C.: 1972, 'Pragmatics'. In D. Davidson and G. Harman (Eds.), *Semantics of Natural Language*, Reidel, Dordrecht, Holland.

Stenning, K.: 1977, 'Articles, quantifiers, and their encoding in textual comprehension'. In R.O. Freedle (Ed.), *Discourse Processes: Advances in Research and Theory, Vol. I: Discourse Production and Comprehension*. Ablex, Norwood, N.J.

Stenning, K.: 1978, 'Anaphora as an approach to pragmatics.' In M. Halle, J. Bresnan, and G.A. Miller (Eds.), *Linguistic Theory and Psychological Reality*, M.I.T. Press, Cambridge, Mass.

Strawson, P.F.: 1950, 'On referring', *Mind* **59**, 320–344.

Vendler, Z.: 1967, 'Singular terms', Chap. 2 of *Linguistics in Philosophy*. Cornell University press, Ithaca, New York, pp. 33–69.

Webber, B.L.: 1978, 'A formal approach to discourse anaphora', Technical Report No. 3761, Bolt, Beranek and Newman Inc., Cambridge, Mass.

Wilson, G.: 1978, 'On definite and indefinite descriptions', *Philosophical Review* **87**, 48–76.