

RIGID DESIGNATION AND REFERENCE

By

SITWAT M. ANSARI, B.A., M.A.

C

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

September 1979

RIGID DESIGNATION AND REFERENCE

DOCTOR OF PHILOSOPHY (1979)
(Philosophy)

McMASTER UNIVERSITY
Hamilton, Ontario

TITLE: Rigid Designation and Reference

AUTHOR: Sitwat M. Ansari, B.A. (University of Delhi, India)
M.A. (University of Delhi, India)
M.A. (McMaster University)

SUPERVISOR: Professor Neil L. Wilson

NUMBER OF PAGES: vi, 186

ABSTRACT

This thesis seeks to provide a theory of reference which preserves the lack of non-rigid uses of proper names and, at the same time, takes into account the social character of naming. The theories of reference which we examine in this thesis either fail to preserve the rigidity of names or ignore the social character of names (usually both) or provide an incorrect principle of decipherment for proper names. Thus, the Description theory and the different versions of the Cluster theory of names suffer from the first two defects, while the Historical Explanation theory and the Causal theory of names provide an incorrect principle of decipherment for referentially ambiguous proper names. The importance of preserving the lack of non-rigid uses of proper names emerges in connection with the problem of trans-world identity. Following Kripke we argue that the problem of trans-world identity which arises in connection with the applied formal semantics for quantified modal logic can be dealt with by means of the notion of a rigid designator. More precisely, we argue that although Quine is correct in thinking that quantified modal logic is committed to some form of essentialism, his failure to come to terms with essentialism is rooted in the fact that he is unable to make sense of the very basic idea of a thing's having certain properties necessarily and independently of the means by which the thing is denoted and the property expressed. However, once we make a distinction between the need to provide a criterion of trans-world identity and the requirement that we make sense of the notion of one and the same object in different counter-factual situations, the problem of trans-world identity admits of an easy solution.

ACKNOWLEDGEMENTS

In the Preface to the Second Edition of The Open Society and its Enemies, Karl Popper states that no book can ever be finished, for while working on it we learn enough to find it immature the moment we turn away from it. I think it is safe to say that Popper's statement applies not just to books but to doctoral dissertations as well. I wish to express my gratitude to Dr. Wilson who enabled me to see the wisdom of Popper's remarks by not only helping me with the earlier versions of this thesis, but by giving me the opportunity of clarifying many problems in detailed discussions which sometimes degenerated into acrimonious philosophical debates. I also wish to express my gratitude to Dr. Radner and Dr. Lehman for their encouragement and to McMaster University for providing me with financial support during the period in which this thesis was written. In the preparation of the final manuscript I have received invaluable assistance from Debbie, and I gratefully acknowledge this. I am especially eager to express my indebtedness to my friends for penetrating criticism and stimulating discussion. Special thanks to Steve, Richard, Dale and Jim.

TABLE OF CONTENTS

	<u>Page</u>
INTRODUCTION	1
Notes and References	13
CHAPTER I <u>QUANTIFYING IN</u>	
Introduction	14
I Quine on Quantifying In: The Problem Stated	15
II Classical Treatments of the Problem	26
1. Frege on Oblique Contexts	26
2. Carnap and the Method of Extension and Intension	30
III Recent Treatments of the Problem	38
1. Smullyan and Scope Distinctions	38
2. Possible-Worlds Semantics	41
Summary	50
Notes and References	51
CHAPTER II <u>RIGID DESIGNATION</u>	
Introduction	54
I The Problem of Trans-World Identity: Frege-Carnap-Kaplan Method	55
II Frege-Russell View	59
1. The Description Theory	59
2. Arguments Against the Frege-Russell View	62
III Semantics of Direct Reference	64
1. Rigid Designation	64
2. Kripke on Proper Names	68
3. Proper Names and Essences	76
IV On Fixing the Reference vs. Supplying a Synonym	83
Summary	86
Notes and References	88
CHAPTER III <u>THE HISTORICAL EXPLANATION THEORY</u>	
Introduction	91
I The Negative Aspect of the Historical Explanation Theory	92
1. Donnellan's Arguments Against the Principle of Identifying Descriptions	92

	<u>Page</u>
2. Donnellan on Attributive and Referential Descriptions	100
3. Humpty Dumpty and Referring	107
IV The Positive Aspect of the Historical Explanation Theory	115
1. Actual Intended Reference	115
2. Termination and Blocks	121
3. A Problem	126
Summary	127
Notes and References	129

CHAPTER IV THE CLUSTER THEORY OF NAMES

Introduction	131
I The Cluster Theory of Names	132
II Toward a Causal Theory of Names	136
1. Kripke's Objections to the Cluster Theory of Names	136
2. The Causal Theory of Names	145
3. Referential Ambiguity	149
III The Causal Theory of Names	153
1. Objection to the Alleged Sufficiency of Condition (I'')	153
2. Intentional Reference	155
3. Termination	158
4. Some Objections	161
IV The Amended Version of the Cluster Theory of Names	164
1. Kripke's Counter-Examples Re-Examined	164
2. Weighting	172
Summary	178
Notes and References	181

<u>BIBLIOGRAPHY</u>	183
-------------------------------	-----

INTRODUCTION

1. The primary aim of this thesis is to present a theory of reference, i.e., to list a set of conditions which have to be satisfied before we can say that a given proper name "N" is a name of such and such an object x. There are various motivations for presenting a theory of reference and these will become clear as we proceed. It is a well known doctrine of Frege and Russell, who represent the classical tradition of modern logic, that a proper name is simply a disguised or abbreviated description, i.e., simply an abbreviation for a phrase of the form "the x such that ϕx ", such as "the teacher of Alexander". According to Frege, the sense of a proper name is given by such a description. Having said that a proper name was an abbreviated description, Frege and Russell had a very simple answer to the question: how is the reference of a name determined on a given occasion of its use by a speaker? If, for example, the name "Aristotle" is simply an abbreviation for "the teacher of Alexander", then the referent of the name "Aristotle" is that person who (uniquely) taught Alexander. The theory that names are abbreviated descriptions also enables us to analyze identity statements. As Kripke points out,

"... we may discover that two names have the same referent, and express this by an identity statement. So, for example ... you see a star in the evening and it's called 'Hesperus' ... We see a star in the morning and call it 'Phosphorus'. Well, then, in fact we find that it's not a star, but is

the planet Venus and that Hesperus and Phosphorus are in fact the same. So we express this by 'Hesperus is Phosphorus'. Here we are certainly not just saying of an object that it's identical with itself. This is something that we discovered. A very natural thing is to say that the real content [is that] the star which we saw in the evening is the star which we saw in the morning ... This, then, gives the real meaning of the identity statement in question; and the analysis in terms of descriptions does this."¹

There are other arguments which can also be regarded as motivations for accepting the Frege-Russell view that names are abbreviated descriptions. One such argument is mentioned by Searle

"We use proper names in existential propositions, e.g., 'there is such a place as Africa', 'Cerberus does not exist'. Here proper names cannot be said to refer, for no such subject of an existential statement can refer. If it did, the precondition of its having a truth value would guarantee its truth, if it were in the affirmative, and its falsity, if it were in the negative ... An existential statement does not refer to an object and state that it exists, rather it expresses a concept and states that that concept is instantiated. Thus, if a proper name occurs in an existential statement it must have some conceptual or descriptive content."²

Now, in spite of the fact that there are compelling arguments for accepting the Frege-Russell view, we argue in the thesis that this theory is, in fact, false. More precisely, we argue that a name is not an abbreviated description and that the reference of a name is not determined via the description associated with the name. Since we reject the Frege-Russell view (i.e., the Description theory), the next step is to inquire about the alternatives to the Frege-Russell view. There are two sets of alternatives to the Frege-Russell view that a name is simply an abbreviated description. Under the first set of alternatives we can include the different versions of the

cluster theory of names and under the second set we can include the historical explanation theory and the causal theory of names. Let us concentrate on the first set of alternatives to the Description theory.

Except for one version of the cluster theory of names (i.e. Wilson's), the other versions of the cluster theory of names (Searle, Strawson, Wittgenstein) differ from the Frege-Russell view only insofar as they maintain that the reference of a name is determined not by a particular definite description but by some cluster or family of descriptions. Thus, for example, in arguing against the Frege-Russell view Searle says,

"Suppose we agree to drop 'Aristotle' and use, say, 'the teacher of Alexander', then it is a necessary truth that the man referred to is Alexander's teacher—but it is a contingent fact that Aristotle even went into pedagogy, though I am suggesting that it is a necessary fact that Aristotle had the logical sum, inclusive disjunction, of properties commonly attributed to him..."³

Furthermore, Searle makes it quite clear that he is using the terms "necessary" and "analytic" interchangeably. Thus, he says,

"Suppose we ask the users of the name 'Aristotle' to state what they regard as certain essential and established facts about him. Their answers would constitute a set of identifying descriptions, and I wish to argue that though no single one of them is analytically true of Aristotle, their disjunction is."⁴

According to Searle, then, a name is not an abbreviation for a single description, but, rather, for a family of descriptions. And the referent of the name is that object which satisfies enough of the cluster of descriptions. Alternatively, we can say that the meaning of the

4

name is equivalent to the meaning of a cluster or family of descriptions and the referent of the name is that object which satisfies enough of the descriptions.

There are strong arguments against the cluster theory of names; for example, it is open to a great many counter-examples. The case against the cluster theory of names has been forcefully argued for by Kripke and Donnellan, and we will deal with these arguments in the thesis. Kripke, however, also argues against both the Frege-Russell view and the cluster theory of names, by arguing for the thesis that names are rigid designators. Here, we will examine this particular argument of Kripke's, for it will enable us to see more clearly one of the motivations for holding the historical explanation theory or the causal theory of names, i.e., for accepting the second set of alternatives to the Frege-Russell view.

2. According to Kripke, proper names are always rigid designators whereas descriptions generally are not rigid designators. By a rigid designator Kripke understands a designator which refers to the same object in all counterfactual situations, namely, the object to which it refers in the actual world. Given this definition of a rigid designator, it follows that descriptions are not always used as rigid designators. For the description "the teacher of Alexander" refers in our world to Aristotle, but in other possible worlds it may be used to refer to any one of a large number of individuals. However, if the meaning of a proper name was equivalent to the meaning of some particular description or even to a cluster of descriptions,

A

then proper names, Kripke wants to argue, would not always be used as rigid designators.

"Suppose the reference of a name is given by a description or a cluster of descriptions. If the name means the same as that description or cluster of descriptions, it will not be a rigid designator. It will not necessarily designate the same object in all possible worlds, since other objects might have had the given properties in other possible worlds, unless (of course) we happened to use essential properties in our description. So suppose we say 'Aristotle is the greatest man who studied under Plato'. If we used that as a definition, the name 'Aristotle' is to mean 'the greatest man who studied under Plato'. Then of course in some other possible world that man might not have studied under Plato and some other man would have been Aristotle."⁵

According to Kripke, then, if we adopt the view that the meaning of a name is equivalent to the meaning of some particular description or cluster of descriptions, we shall be unable to account for uses of names in the description of counterfactual situations. Furthermore, since it is the contention of a cluster theorist like Searle that names are analytically tied to a cluster of descriptions, it follows that the cluster theorist will be unable to account for the uses of names in the description of counterfactual situations.

At first glance it appears that there is a way of circumventing Kripke's argument that if the meaning of a name was given by some particular description or cluster of descriptions then the name will not necessarily designate the same object in all possible worlds, i.e., the name will not be a rigid designator. Consider, for example, the following sentence

(1) Aristotle might not have taught Alexander.

Let us assume, furthermore, that the name "Aristotle" is equivalent

in meaning to the description "the teacher of Alexander". Substituting the phrase "the teacher of Alexander" for the name "Aristotle" in (1), we get

(2) the teacher of Alexander might not have taught Alexander.

If we abbreviate "x is the teacher of Alexander" as "Tx", then we can write (2) as

(3) $\text{pos} \sim T(\exists x)(Tx)$.

The problem with (2), though, is that it can be understood in two ways. It could be read either as

(4) It might have been the case that the teacher of Alexander did not teach Alexander.

or as

(5) the teacher of Alexander might not have taught Alexander.

Where (4) and (5) are to be written, respectively, as

(6) $\text{pos} [(\exists x)(Tx)] \sim T(\exists x)(Tx)$

and

(7) $[(\exists x)(Tx)] \text{pos} \sim T(\exists x)(Tx)$

The sentences (6) and (7) in turn expand as

(8) $\text{pos} (\exists y)[\sim Ty \ \& \ (x)(Tx \equiv x = y)]$

and

(9) $(\exists y) [\text{pos} \sim Ty \ \& \ (x)(Tx \equiv x = y)]$

Now, (8) expresses the sense in which (2) is false, i.e., the sense in which we deny it by saying that it could not have been true that the teacher of Alexander did not teach Alexander. However, (9) expresses the sense in which (2) is true. In (9) we are treating the definite description as having, as its constant referent, that referent

which it has in the actual world, and this amounts to taking the description as a rigid designator, i.e., taking the scope of the description to lie outside the scope of the modal operator. It seems, then, that it is possible to construe a name as a rigid designator even in those cases where the meaning of the name is equivalent to the meaning of some particular description, since the description which is substituted for the name satisfies the initial characterization, i.e., it designates the same object in all possible worlds. Are we then to conclude that Kripke is mistaken in thinking that if a name was analytically tied to some particular description it would not necessarily designate the same object in all possible worlds? I don't think so for the following reason.

To start with, if we assume that a definite description is being used non-rigidly, then it is important to note that the way in which the reference of the description is determined in the actual world is carried over into each particular possible world. In any possible world, the referent of "the man who taught Alexander" is the one and only one object (if any) which, in that world, satisfies the open sentence "x is a man who taught Alexander". The situation is different with regard to a designator used rigidly. To say that there is a possible world in which Aristotle did not teach Alexander, implies (in the appropriate sense) that, in that world, the reference of the name "Aristotle" is not determined in the way in which it is determined in the actual world; rather it is determined in that world via its reference in the actual world, i.e., as being to the same man who is its referent in the actual world. Alternatively, we can say

that to assign to a term a reference varying from one possible world to another is just to take it as having, in each world, the description (if any) which it has in that world; conversely, to assign it a constant reference, is to take it as having, in each world, just that reference which it has in the actual world. But to take a description in the former of these two ways is precisely to treat it as being within the scope of the modal operator while to take it in the second way is to treat it as falling outside the scope of the modal operator.

Now, as we have seen, modal sentences such as (2) will have two non-equivalent readings depending upon whether the description "the teacher of Alexander", is being used rigidly or non-rigidly, i.e., as (7) or as (6). However, Kripke wants to maintain that proper names are always used as rigid designators. But if the name "Aristotle" was equivalent in meaning to the description "the teacher of Alexander", it would have at least some uses which were non-rigid, viz., (6). Therefore, we may conclude that a name is not equivalent in meaning to a particular description, and that Kripke is right in thinking that if a name was analytically tied to some particular description it would not necessarily designate the same object in all possible worlds. Since the above argument (with obvious modifications) can be applied equally well to the theory that the meaning of a name is equivalent to the meaning of a cluster or family of descriptions, it follows that a name is not equivalent in meaning to a cluster of descriptions. We have, of course, been assuming thus far that the thesis that proper names are always used rigidly is a correct thesis. We argue for this in the main text.

2. We are now in a position to see more clearly one of the motivations for holding the causal theory of names or the historical explanation theory of names, i.e., for accepting the second set of alternatives to the Frege-Russell view. If the reference of a name is not determined via a cluster of descriptions which are analytically tied to the name, then we are faced with the following question: how is the reference of a name determined on a particular occasion of its use by a speaker? The causal theorist and the historical explanation theorist give us an answer to this question. According to them, the reference of a name on a particular occasion of its use by a speaker is determined via a three-place relation which holds between the speaker, the name and the object named, and which does not consist in the object named satisfying the cluster of descriptions which the speaker associates with his use of the name. A rough statement of the theory would be the following:

"Someone, let's say, a baby, is born; his parents call him by a certain name. They talk about him to their friends. Other people meet him. Through various sorts of talk the name is spread from link to link as if by a chain. A speaker who is on the far end of this chain, who has heard about, say Richard Feynman, in the market place or elsewhere, may be referring to Richard Feynman even though he can't remember from whom he ever heard of Feynman. He knows that Feynman was a famous physicist. A certain passage of communication reaching ultimately to the man himself does reach the speaker. He then is referring to Feynman even though he can't identify him uniquely. He doesn't know what a Feynman diagram is, he doesn't know what the Feynman theory of pair production and annihilation is. Not only that: he'd have trouble distinguishing between Gell-Mann and Feynman. So he doesn't have to know these things, but, instead, a chain of communication going back to Feynman himself has been established, by virtue of his membership in a community which passed the name on from link to link ..."⁶

So the reference of a name, according to Kripke, is determined by means of a chain of communication which holds between the speaker's use of a name and the object which was initially bestowed with the name. There are no important differences between the views of Kripke and Donnellan. The only difference seems to reside in the way in which the historical chain of communication and the causal chain of communication terminate at a given object. For our purposes this difference can be more or less ignored, although it will become clear in the thesis. Of greater importance is the fact that the causal theory of names and the historical explanation theory of names preserve the rigidity of names, or, to be more precise, preserve the lack of non-rigid uses of a name. For since the referent of the name is not determined via the cluster of descriptions associated with the name, the referent of the name is going to be the same in all possible worlds, i.e., the name is going to rigidly designate the same object in all possible worlds. So if the causal theory of names and the historical explanation theory of names were correct they would give a nice picture of the way in which proper names function in the description of counterfactual situations. Unfortunately, there are difficulties with both theories. The difficulties revolve around the principle which these two theories provide for the decipherment of referentially ambiguous names. We discuss them in the third and fourth chapters. The amended version of the cluster theory of names which we offer is meant to overcome the difficulties facing the cluster theory of names and is also meant to preserve the lack of non-rigid uses of proper names.

4. The unamended version of the cluster theory of names, as we have seen, is faced with two serious difficulties. The first is that a proper name will have certain non-rigid uses, since, on the cluster theory of names, a name is analytically tied to a cluster of descriptions. The theory presented by us preserves the lack of non-rigid uses of a name. It does this by asserting that the reference of a name on a particular occasion of its use by a speaker is determined via a cluster of descriptions which fix the reference of the name but are not equivalent to it in meaning. Consider, for example, the name "Aristotle". If the cluster of descriptions (both attributive and definite descriptions) are only used to fix the reference of the name "Aristotle", then the referent of the name will be the same in all possible worlds. Thus, when we say counterfactually, "suppose Aristotle had not gone into philosophy", we need not mean "suppose the student of Plato, the teacher of Alexander, etc., had not gone into philosophy"; we need only mean "suppose that that man had not gone into philosophy", i.e., the man who is the referent of the name "Aristotle" in the actual world.⁷

The second difficulty faced by the unamended version of the cluster theory of names has to do with the counter-examples to the theory. The amended version of the cluster theory of names provides intuitively correct responses to Kripke's counter-examples to the unamended version of the cluster theory of names, while at the same time accounting for Kripke's criticism of the unamended cluster theory, namely, that it ignores the social character of naming.

5. Once we have a theory of reference which preserves the lack of non-rigid uses of proper names we can deal with the problem of trans-world identity which arises in connection with the applied formal semantics for quantified modal logic. To make sense of modal sentences such as " $(\exists x) \text{ nec } Fx$ " we need some theory of individuation which accounts for identification of objects from one possible world to another. Suppose, for example, that there is a possible world H (distinct from the actual world) in which Aristotle is not the tutor of Alexander. In the world H he may also lack other properties which he has in this world. But then how can we possibly identify Aristotle in the world H? It cannot be argued that we identify Aristotle in the world H by means of the properties which he possesses in the actual world. For it could well turn out that in the world H someone else possesses the properties which we use to characterize Aristotle in this world. We argue in the thesis that the notion of a rigid designator can be used to solve the problem of trans-world identity. This is basically Kripke's argument. However, we point out that the notion of a rigid designator can be used to solve the problem of trans-world identity because it picks out the same individual across possible worlds by means of its individual essence. And that, therefore, Quine is right in claiming that quantified modal logic is committed to essentialism.

FOOTNOTES

¹Kripke, S. "Naming and Necessity" in Semantics of Natural Language, Davidson, D. and Harman, G. (eds.) (D. Reidel Publishing Company: Dordrecht, 1972), pp. 256-257.

²Searle, J. "The Problem of proper names" in Semantics, Steinberg, D. and Jakobovits, A. (eds.) (Cambridge University Press; London, 1971), pp. 135-136.

³Ibid., p. 140.

⁴Ibid., p. 138.

⁵Kripke, S. "Naming and Necessity", Semantics of Natural Language, op. cit., p. 276.

⁶Ibid., pp. 298-299.

⁷The argument is basically Kripke's. See, for example, "Naming and Necessity", op. cit., p. 276.

CHAPTER 1

QUANTIFYING IN

INTRODUCTION: Faced with the success of the various systems of quantified modal logic constructed by Carnap, Kanger, Kripke and others, one might think that the insistence of Quine that the interpretation of quantified modal logic was beset with difficulties which might even force a collapse of modal distinctions, was antiquated. Quine, however, has not altered his views on quantified modal logic ever since he first explicitly formulated them in 1941.¹ The aim of this chapter is to examine Quine's attack on modal logic and to determine whether Quine is justified in clinging so tenaciously to his former views. It will be argued in this chapter that he is. More specifically, since Quine is concerned with the problem of interpreting modal logic, it will be argued that although semantical considerations on modal logic show that modal logic can be given an interpretation which appears to be offhand intelligible, the interpretation is in fact philosophically dubious. In other words, we will argue that the semantical account of quantified modal logic (i.e., the possible world account) does not suffice to make it philosophically respectable. The main relevance of the semantical account seems to be that it shows that if one is to take exception to quantified modal logic it has to be on philosophic grounds, and not on logical ones. This chapter is divided into three sections. In the first section we will state Quine's objections to the problem of

interpreting modal logic. And in the second and third sections we will state and examine two attempts to render quantification into opaque contexts intelligible, i.e., the attempts by Frege and Smullyan. It will be argued that none of these attempts succeed but that the best hope for modal logic is in accepting Smullyan's solution to the difficulties raised by Quine.

I

QUINE ON QUANTIFYING IN: THE PROBLEM STATED

Quine's criticism of modal logic is directed, specifically, against the third and gravest degree to which we may allow our logic and semantics to embrace the idea of necessity.² The first degree of modal involvement is expressed by means of a semantical predicate which is attached to names of statements. Thus, we can attach the predicate "Nec" to a noun or a singular term which is a name of a statement. For example, we can write

(1) Nec "9 > 5"

The semantical predicate "Nec" applies to a name to yield a statement. When used as a semantical predicate Nec "p" is interpreted as being true if and only if "p" is logically true. Thus we will have formulae such as "Nec (p v ¬p)", "Nec (p ⊃ q) ⊃ (Nec p ⊃ Nec q)", etc. This use of "Nec" is generally held to be unobjectionable. It is important to note, however, that the letters "p", "q", etc. are not schematic letters which depict the form of actual statements but special variables which replace names of statements.

The second grade of modal involvement, is expressed by attaching

the logical operator "nec" to statements to form other statements. So long as necessity is adopted in the form of a statement operator it can be converted into a semantical predicate which attaches to names of statements. Thus, for example, we can convert the statement operator "nec" in the statement

(2) $\text{nec } (9 > 7)$

into the semantical predicate "Nec" merely by supplying quotation marks in the following manner

(3) $\text{Nec } "9 > 7"$

One reason for translating the statement operator "nec" into its semantical counterpart "Nec" is that it serves to remind us that the context in question is referentially opaque. A referentially opaque context is one where a term does not simply refer to its object. Consider, for example, the following statement:

(4) "Tully" contains six letters.

The statement (4) says nothing about the man Tully and the occurrence of the term "Tully" in (4) is not referential. The same is true of the following statement:

(5) "9 > 5" contains three characters.

The statement (5) says nothing about the numbers 9 and 5. The principle of the substitutivity of identity can be used to determine whether a term occurs referentially in a given context. Since

(6) $9 = \text{the number of planets}$

and since whatever is true of 9 is true of the number of planets, we should be able to substitute the term "the number of planets" for "9" in (5) and come out with a truth. But since the substitution of "the

number of planets" for "9" in (5) leads to the falsehood

(7) "the number of planets > 5" contains three characters.

we can say that the position occupied by the term "9" in (5) was not referential. Thus, following Quine, we can say that a context is referentially opaque when "by putting a statement ϕ into that context we can cause a purely referential occurrence in ϕ to be not purely referential in the whole context."³

Thus far we have been talking about the second degree of modal involvement. However, once we turn to the third degree of modal involvement and start using "nec" as a sentence operator in predicate logic we can no longer translate it into the semantical predicate "Nec". We cannot, for example, write

(8) nec ($x > 7$)

as

(9) Nec " $(x > 7)$ "

"'Nec' has been understood up to now as a predicate true only of statements, whereas (9) attributes it rather to an open sentence and is thus trivially false, at least pending some deliberate extension of usage. More important, whereas (8) is an open sentence with free 'x', (9) has no corresponding generality; (9) is simply a statement about a specific open sentence. For, it must be remembered that ' $x > 5$ ' in quotation marks is a name of the specific quoted expression, with fixed letter 'x'. The 'x' in (9) cannot be reached by a quantifier."⁴

Just because we cannot quantify into a quotation context so it wouldn't make any sense to write (9) as

(10) $(\exists x)(\text{Nec } 'x > 7')$.

Since " $x > 7$ " is a name of the expression occurring in the quotation marks, the "x" in " $x > 7$ " is a fixed letter. It cannot be reached by

a quantifier. The trouble is that in modal logic one wants to quantify into necessity contexts and we cannot quantify into quotations. So it will not do to treat "nec" as a sentence operator, as a mere shorthand for the semantical predicate "Nec" and a pair of quotation marks. The problem we are faced with is the following. Either we regard "nec" as a sentence operator, as a mere shorthand for the semantical predicate "Nec" and a pair of quotation marks in which case we are not allowed to quantify into those contexts; or else we try and make sense of sentences of the form

$$(11) (\exists x) \text{ nec } (x > 7),$$

The stumbling block to the second alternative is that there are numerous difficulties which can be raised concerning the sense of (11). These difficulties were first raised, as far as I know, by Quine in his "Notes on existence and necessity".⁵ In that article Quine asked "... would 9, that is, the number of planets, be one of the numbers necessarily greater than 7?"⁶ And he pointed out that such an affirmation would be true in the form (1) and false in the form

$$(12) \text{ nec } (\text{the number of planets} > 7)$$

and yet we are confronted with the truth of

$$(6) 9 = \text{the number of planets}$$

Unlike (6), the statement

$$(13) \text{ nec } (\exists x)(x > 7)$$

that is, the statement "Necessarily something is greater than 7" makes sense, according to Quine, because it is in fact a true statement. The difference between (13) and (11) is the difference between quantification within an opaque context and quantification into an opaque context.

Consider, for example, the following sentences: "It is possible that, for some x , x will never cease to exist" and "For some x , it is possible that x will never cease to exist". The former involves only quantification within an opaque context and means that it is logically possible (given that we are operating with the logical sense of possibility) that there should be something which will never pass out of existence; the latter involves quantification into an opaque context, and means that, among the objects which there actually are, there is at least one of which it is logically possible that it should never pass out of existence.

In the same paper ("Notes on existence and necessity"), Quine also pointed out the failure of the law of substitutivity of identity in modal (alethic) contexts, such as "necessarily ..." at least when necessity is understood in the sense of analyticity. He pointed out that from (1) and (6) we cannot infer (12), with the help of the law of substitutivity of identity. Quine's conclusion was that necessity and possibility contexts are similar to quotation contexts which do not admit pronouns which refer to quantifiers anterior to the context.

In his review of Quine's "Notes on existence and necessity", Church agreed with Quine that modal contexts are opaque, but he insisted that this does not prevent variables within the modal contexts from being bound by a quantifier anterior to the context, provided that the quantifier has an intensional range, for instance, a range of attributes.⁷ Quine replied to this in a much later article and we will come to that, but in the meantime Quine was able to sharpen his criticism of quantification into modal contexts.

Thus, in another article published in 1947, much later than his "Notes on existence and necessity" Quine once again took up the question of referential opacity.⁸ In this article he argued that acceptance of quantified modal logic leads us to accept the view that there are only a multitude of distinguishable entities and no concrete objects. To show this Quine uses the symbol "C" to express the relation of "congruence" which Venus, the morning star and the evening star bear to one another and to themselves. Then we have

(14) Morning Star C Evening Star & nec (Morning Star C Morning Star).

Quine also adopts a criterion of existential generalization according to which an existential generalization holds if there is a constant whose substitution for the variable in the open sentence would make the open sentence true. So, from (14) with the help of the criterion we have

(15) $(\exists x)(x \text{ C Evening Star} \ \& \ \text{nec } (x \text{ C Morning Star}))$.

But we also have

(16) Evening Star C Evening Star & -nec (Evening Star C Morning Star).

Again, applying the criterion to (16) we get

(17) $(\exists x)(x \text{ C Evening Star} \ \& \ \text{-nec } (x \text{ C Morning Star}))$.

Since (14) and (16) are mutual contraries, it can't be the same individual which satisfies the operands of (15) and (17). Quine concludes that there are at least two objects congruent to Evening Star. According to Quine, a similar argument can be constructed by the introduction of the term "Venus" to show that a third such object can be inferred. Quine suggests that there are two options open to the modal logician who finds this repudiation of concrete objects from his domain of

discourse uncongenial. The first alternative is to regard modal logic as a proper part of the total logic to which the modal logician is prepared to subscribe so that the ontology of "individual concepts" will come to be regarded only as a part of the total ontology which embraces material objects and maybe even classes. In this case the variables of modal logic would have to be distinguished from those of the total logic so that in the total logic we would not be able to arbitrarily apply modal operators to any matrix and then quantify with respect to the variable occurring in the matrix. The second alternative for the modal logician is to insist on the universal applicability of his quantified modal logic but to repudiate the partial criterion of existential generalization which was offered by Quine. However, in this case it is not clear what other criterion can be offered in its place.

In his article "Reference and modality" Quine brought together the arguments he had put forth in "Notes on existence and necessity" and "The problem of interpreting modal logic", but he also added new arguments to these to the effect that we cannot quantify into modal contexts.⁹ In this article Quine tries to show the meaninglessness of quantification into modal contexts without reverting to an "expository interplay between singular terms":

"Whatever is greater than 7 is a number, and any given number x greater than 7 can be uniquely determined by any of various conditions, some of which have ' $x > 7$ ' as a necessary consequence and some of which do not. One and the same number x is uniquely determined by the condition:

$$(32) \quad x = \sqrt{x} + \sqrt{x} + \sqrt{x} + \sqrt{x}$$

and by the condition:

(33) There are exactly x planets,

but (32) has ' $x > 7$ ' as a necessary consequence while (33) does not. Necessary greatness than 7 makes no sense as applied to a number x ; necessity attaches only to the connection between ' $x > 7$ ' and the particular method (32), as opposed to (33), of specifying x ."¹⁰

Similarly Quine argues that

(18) $(\exists x)$ (necessarily if there is life on the Evening Star then there is life on x).

is meaningless because the physical object which fulfills the condition

(19) necessarily, if there is life on the Evening Star then there is life on x

can be uniquely determined by different conditions not all of which have (19) as a necessary consequence.

"Necessary fulfilment of (19) makes no sense as applied to a physical object x ; necessity attaches, at best, only to the connection between (19) and one or another particular means of specifying x ."¹¹

In Word and Object, Quine considers a tentative solution to this difficulty.¹² The solution consists in narrowing the universe of objects

so as to exclude those objects which can be specified in different ways that are not necessarily equivalent. He now considers admitting only those objects which satisfy the following condition:

(C) If one and the same object is uniquely determined by two open sentences then the two sentences are necessarily equivalent.

Condition (C) has the consequence that modal distinctions collapse.

Suppose we write (C) schematically as

(20) $[(w)(Fw \equiv w = x) \ \& \ (w)(Gw \equiv w = x)] \supset \text{nec}(w)(Fw \equiv Gw)$

From (20) we can deduce 'necessarily p ' where ' p ' is any true sentence.

Let y be any object in our purified universe of discourse and let $x = y$.

Then we have

$$(21) \quad (w)[(p \ \& \ w = y) \equiv w = x]$$

$$(22) \quad (w)(w = y \equiv w = x)$$

Next, by substituting ' $p \ \& \ \textcircled{1} = y$ ' for ' $F\textcircled{1}$ ' and ' $\textcircled{1} = y$ ' for ' $G\textcircled{1}$ ' in (20) we get

$$(23) \quad \{(w)[(p \ \& \ w = y) \equiv w = x] \ \& \ (w)(w = y \equiv w = x)\} \supset \text{nec} \\ (w)[(p \ \& \ w = y) \equiv w = y]$$

From (23) by (21) and (22) we get

$$(24) \quad \text{nec} \ (w)[(p \ \& \ w = y) \equiv w = y]$$

By instantiation and necessity elimination, we have

$$(25) \quad (p \ \& \ y = y) \equiv y = y$$

(25) implies 'p'. Hence, since what is implied by a necessary truth is a necessary truth, (24) implies that necessarily p.

We see, then, that narrowing our domain of objects will not help since it only leads to a collapse of modal distinctions. Since nothing is assumed here about the objects over which we are quantifying, restricting the values of one's variables to intensional objects will not help in averting this collapse of modal distinctions. In the article "Reply to Professor Marcus" Quine points out that Church was wrong in his review of "Notes on existence and necessity" when he said that intensionalizing the values of variables would prevent two specifications of the same object which only coincide contingently.¹³

"Now on this latter point Church was wrong. I have been slow to see it but the proof is simple. Anything x, even an intension, is specifiable in contingently coincidental ways if specifiable at all. For suppose x is determined uniquely by the condition ' ϕx '. Then it is also determined uniquely by the conjunctive condition ' $p \ \& \ \phi x$ ' where 'p' is any

truth, however irrelevant. Take 'p' as an arbitrary truth not implied by ' ϕx ', and these two specifications of x are seen to be contingently coincident: ' ϕx ' and 'p. ϕx '.¹⁴

In "Three Grades of Modal Involvement", Quine draws one further consequence of quantifying into modal contexts.¹⁵ This is what he calls "Aristotelian essentialism", and by which he understands the doctrine that some of the attributes of a thing are essential to it, necessary of the thing regardless of the way in which we refer to it. For example, a man is essentially rational, not merely qua man, but qua himself. According to Quine, this means that we prefer certain ways of specifying (uniquely) an object x while regarding other ways of uniquely specifying the object x as uncongenial. The ways we somehow favour are supposed to better reveal the "essence" of the object x. Thus, for example,

$$(26) \quad x = \sqrt{x} + \sqrt{x} + \sqrt{x} \neq \sqrt{x}$$

is to be preferred to

$$(27) \quad \text{There are exactly } x \text{ planets}$$

since consequences of (26) can be looked upon as necessarily true of the object which is 9, whereas consequences of (27) are only contingently true of it. Quine's conclusion is "so much the worse for quantified modal logic."¹⁶ We will examine Quine's objection to quantified modal logic, i.e., that it commits us to some form of what he calls 'Aristotelian essentialism' in the next chapter. It will be seen that quantified modal logic is indeed committed to essentialism, to be more precise, to the notion of an individual essence. However, as it will become clear later on, we do not have to accept Quine's conclusion:

"so much the worse for quantified modal logic". But apart from Quine's objection to essentialism, the rest of the difficulties raised by Quine regarding quantification into modal contexts all revolve around the failure of the law of substitutivity of identity in modal contexts. We can sum up the difficulty regarding the failure of the law of substitutivity of identity in modal contexts in the following way. Consider the following two sentences:

(1) $\text{nec } (9 > 7)$

(6) The number of planets = 9

Now, according to Quine, one of the fundamental laws governing the notion of identity is the substitutivity of identity which states that given any true identity statement, we can substitute either term of that identity statement for the other in any other true statement without altering the truth-value of the latter statement. However, the substitution of the expression "the number of planets" for the term "9" in the true statement (1) gives us the obviously false statement

(12) $\text{nec } (\text{the number of planets} > 7)$

Something has gone wrong. We have deduced a false conclusion from true premisses using what seem to be perfectly legitimate modes of reasoning.

In the next section we will state and examine two attempts to deal with the paradox that arises as a result of the failure of the law of substitutivity of identity in modal contexts, namely, the attempts by Frege and Carnap.

II

CLASSICAL TREATMENTS OF THE PROBLEM.

1. FREGE ON OBLIQUE CONTEXTS: Since contexts which Quine calls "opaque" occur frequently in substantial clauses, clauses governed by "that" or "whether", Frege refers to the topic as that of "indirect speech" (oratio obliqua). It is a well-known doctrine of Frege's that proper names occurring in a referentially opaque context cannot have their ordinary reference.¹⁷ His reasons could be summarized in the following way. Consider, for example, the following statement

(28) Frege believed that Kepler died in misery.

Now, suppose in this case the proper name "Kepler" in the subordinate clause stood for its customary referent, namely, the man Kepler. In that case we would have to take the sentence appearing in the subordinate clause as saying something about the man Kepler. We would then have to take

(29) Frege believed that x died in misery

as a predicate true of the man Kepler. However, if a predicate is true of an object it is true of the object no matter how we refer to the object. Suppose now, that we choose to refer to the man Kepler by means of the name "the discoverer of the elliptic form of the planetary orbits". Since the name "the discoverer of the elliptic form of the planetary orbits" refers to the same man as the name "Kepler", the statement

(30) Frege believed that the discoverer of the elliptic form of the planetary orbits died in misery

should have the same truth-value as the statement (28). But clearly it doesn't. Therefore, the name "Kepler" does not refer to the man Kepler in (28). Arguments similar to this can be constructed to apply to contexts governed by alethic modalities.

Given that a proper name cannot stand for its customary reference in an opaque context we have to ask ourselves what it could stand for. To begin with, we may note that the above argument may be generalized. The same considerations that apply, for Frege, to proper names apply equally to declarative sentences which occur in contexts governed by substantival chances or modal (alethic) operators. Since, for Frege, the reference of a sentence is the truth-value, the replacement of one part of a sentence by another having the same reference as the one replaced will not alter the truth-value of the original sentence although it may alter its sense.¹⁸ The substitution, however, may not work in the case of an opaque context. More specifically, the replacement of a sentence governed by a substantival clause by another with the same truth-value may not preserve the truth-value of the original complete sentence. Now, in order to answer our former question, namely, what do expressions in an opaque context stand for, we must ask what replacements would leave the truth-value of a sentence like (27) unaltered. Frege's answer was that the truth-value of a sentence in a subordinate clause would be unaltered if we replaced it by another sentence with the same sense:

"In such cases it is not permissible to replace one expression in the subordinate clause by another having the same customary reference, but only by one having the same indirect reference, i.e., the same customary sense."¹⁹

Normally, we use a sentence to express a sense (thought) but in doing this we do not talk about the sense. However, when we say, for example that "John said that Scott was the author of Waverly" what we are talking about is the thought expressed by John, about the sense of the words he used.

Frege's main idea is that there are no real opaque occurrences. The appearance of opacity created by apparent failures of substitutivity and the like is due to confusion about what the reference of a given expression is. It is easy to see how Frege would have handled examples like (1), (8) and (7). He would say that it does not follow from the truth of (8) that the name "the number of planets" has the same reference in (7) that it does in (8).

According to Frege, then, names in oblique contexts acquire new nominata and new senses. The nominatum of a name in an oblique context is its sense in an ordinary context. Thus, two names in an oblique context can be interchanged freely salva veritate if their sense in an ordinary context is the same. But Frege's method turns out to be much more complicated than it appears at first sight. On Frege's own principles it is sense which could be said to determine reference and not the other way round, reference does not determine sense. We have seen that the referent of a name in an oblique context is its customary sense. What, however, is the sense of a name in an oblique context? Reference does not determine sense. So just knowing what the referent of a name in an oblique context is, is not enough to enable us to determine what its sense in an oblique context is. Moreover, it is doubtful that we will ever be able to say what the sense of an expression

when it occurs in an opaque context is to be taken to be. Consider, for example, the name "Hesperus". One might say that since "Hesperus", when it occurs in opaque contexts usually stands for what in transparent contexts is its ordinary sense, in opaque contexts its sense is the same as the sense of that expression which in transparent contexts stands for the customary sense of "Hesperus". That is, the indirect sense of "Hesperus" in opaque contexts is the same as the customary sense of "the sense of 'Hesperus'." This, however, doesn't work. It means for one thing, that the doubly oblique sense of "Hesperus" would be the same as the sense of the expression "the sense of ..." applied twice to the name "Hesperus". It is not clear how repeated applications of the expression (or operator if you like) "the sense of ..." are supposed to differ from just one application of it to the name "Hesperus". It is not even clear how the sense of "the sense of 'Hesperus'" is different from the sense of "Hesperus".

Now, consider a statement containing two opacity-inducing clauses such as

(30) Jones said that Jack wanted to know whether Sikorski was bald.

Since the name "Sikorski" occurs within the clause following "said that ...", on Frege's theory it must stand for the customary sense it would have in the statement "Jack wanted to know whether Sikorski was bald". However, this statement also contains an opacity-inducing operator, namely, "wanted to know whether".²⁰ So combining the two opaque contexts we can say that the name "Sikorski" in (30) has a doubly oblique reference (and a doubly oblique sense). But it was argued above that we cannot say what the sense (simple) of an expression

in an oblique context is. Then, since according to Frege the sense of an expression determines its reference, it follows that we cannot even say what the reference (simple) of the name "Sikorski", i.e., its customary sense, in the statement (30) is. But if we do not know what its customary sense is then how can we determine its reference? It follows that we do not know how to assess the truth-value of a statement in which two opacity-inducing operators occur. Frege must have gone wrong somewhere.

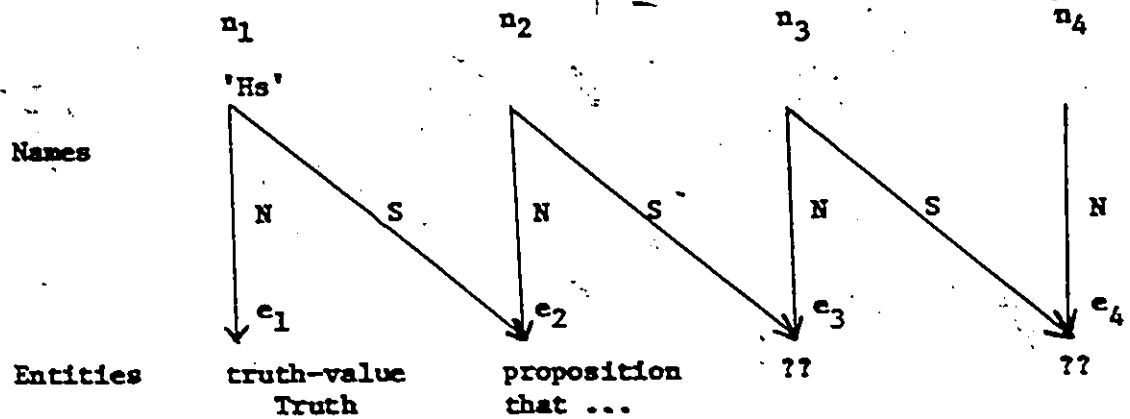
Since Frege's theory leads to such unintuitive results we have to ask ourselves whether or not it can be modified so as to restore the original simplicity of the theory. Thus, for example, Kaplan writes:

"My own view is that Frege's explanation by way of ambiguity, of what appears to be the logically deviant behaviour of terms in intermediate contexts is so theoretically satisfying that if we have not yet discovered or satisfactorily grasped the peculiar intermediate objects in question then we should simply continue looking."²¹

By "intermediate contexts" Kaplan means opaque or oblique contexts and the "intermediate objects" are the nominatum of expression in these contexts. Now, Carnap's method of extension and intension can be regarded as a modification of Frege's theory and we will examine it next.

2. CARNAP AND THE METHOD OF EXTENSION AND INTENSION: According to Carnap, Frege's method of dealing with oblique contexts leads to a multiplication of entities. Carnap, in his Meaning and Necessity, illustrates this schematically by means of the following diagram.²²

In the diagram 'Hs' is the sentence 'Scott is human'. An arrow with 'S' stands for the sense relation.



According to Frege, the entity named by 'Hs' is the truth-value of 'Hs'. In the diagram this is the entity e_1 . However, there is also the proposition that Scott is Human which is the sense of 'Hs'. Thus, if we wish to speak about this proposition, i.e., the entity e_2 we need another name, distinct from n_1 . Now, if we let n_2 be the name of e_2 then n_2 will also have a sense for which we will need another name and so on. Thus, Frege's method leads to an infinite number of entities, and in order to speak about these entities we require an infinite number of names. For his own method Carnap claims the advantage of requiring only one expression in the object language. Thus, he says

"On the basis of the method of extension and intension, on the other hand, we need in the object language, instead of an infinite sequence of expressions, only one expression ... and we speak in the metalanguage only of two entities in connection with the one expression, namely, its extension and its intension (and even these are, as we shall see later, merely alternative ways of saying the same thing)."²³

Let us examine these claims. We need an exposition of the method of extension and intension. In what follows capital letters A, B, C ... function as metalinguistic variables which take expressions in the

object language as values. Logical constants are to be regarded as names of themselves when combined with capital letters; and a string of names of expressions is the name of the expression formed by concatenating the named expression in the order named. Moreover, in the informal exposition which follows we will only be concerned with individual expressions (constants or descriptions), sentences and predicates.

The two fundamental concepts employed by Carnap in Meaning and Necessity are truth and L-truth. A sentence A_m is L-true (in a semantic system S) if and only if A_m holds in every state-description (in the system S). A state-description is a maximally consistent set of sentences, i.e., a set of sentences which contains for every atomic sentence either this sentence or its negation, but not both, and no other sentences. If A_m and A_n are two expressions of the same type (individual expressions, predicates with the same number of places and sentences) then A_m and A_n are said to be equivalent if and only if $A_m \equiv A_n$ is true, and L-equivalent if and only if $A_m \equiv A_n$ is L-true.²⁴ A context A_m is called extensional if every expression A_n which is obtained by replacing an expression in A_m by an equivalent expression is equivalent to A_m . A context A_m is said to be intensional with respect to the occurrence of A_n in A_m if it does not contain any extensional context and the occurrence of A_n in A_m can be replaced with any other expression L-equivalent to A_n . Carnap then calls any context intensional if it contains at least one context which is intensional with respect to the occurrence of an expression and whose other contexts are extensional. In an extensional context equivalent and L-equivalent

expressions can be substituted *salva veritate*. In an intensional context only L-equivalent expressions may be substituted *salva veritate*.

It is now apparent how Carnap's method may be applied to the problem of quantifying into opaque contexts. The failure of the true sentence (1) to remain true when the expression "the number of planets" is substituted for the expression "9" is explained by pointing out that (1) is an intensional context, while "the number of planets" and "9" are, although equivalent, not L-equivalent.

Thus far we have seen how with the help of the notions of truth and L-truth Carnap was able to talk about expressions of the object language. Before we evaluate Carnap's method we have to show how Carnap was able to translate sentences of the object language into the meta-language. To give examples we will confine our attention to the metalanguage M^1 , the extensional object language S_1 , and the intensional object-language S_2 .²⁵ The meta-language M^1 is meant to be neutral. It does not contain the phrases 'the property Human' and 'the class Human' but only the neutral expression 'Human'. Let 'RA' mean 'is a rational animal' in S_1 and let 'H' mean 'is human'. A characteristic sentence of S_1 would then be

$$(31) \quad RA \equiv H$$

(31) is a L-true sentence of S_1 . Since S_1 is an extensional object language we cannot express the analytic character of (31) in S_1 but we can do so in M^1 by pointing out that 'RA' and 'H' are L-equivalent in S_1 . However, the analytic character of (31) can be expressed in S_2 because S_2 contains the modal operator 'N' (read 'necessarily'). 'N' is defined in terms of the notion of state-description in such a

way that a sentence of S_2 such as $N(\phi)$ will be true if and only if, ' ϕ ' is L-true. Thus, we express the analytic character of (31) in S_2 by writing

(32) $N(RA \equiv H)$.

Furthermore, (32) may be translated into the meta-language M^1 as

(33) Rational Animal is L- equivalent to Human.

The next question which arises is what entities are presupposed in order to interpret the sentences of Carnap's languages.

To each designator there corresponds exactly two entities. The entities differ depending on the type of designator in question. We can represent the situation schematically in the following way:

	Extension	Intension
Individual expressions	Individuals	Individual Concepts
Predicates	Classes	Properties
Sentences	Truth-Value	Proposition

According to Carnap no two distinct individual constants correspond to the same entity. Individual descriptions may correspond to the same entity as individual constants. Moreover, two L- equivalent individual descriptions may correspond to the same entity. There is one entity corresponding to each sentence and one entity corresponding to each predicate. The relation between an entity and a designator is expressed by means of "designates" and "L- designates"; and these notions are part of the meta-language M^1 .²⁶ Entities are equivalent if the corresponding designators are equivalent and L- equivalent if the corresponding designators are L- equivalent.

The difficulty with Carnap's theory is that we are left with

intensions. This has already been pointed out by Davidson²⁷, and Quine²⁸. It can be illustrated by means of individual expressions although the same reasoning applies to predicates and sentences.

Consider the following sentences

(34) Scott is equivalent to the author of Waverly

and

(35) Scott is equivalent to Scott.

Although these two sentences are true sentences of M^1 , (35) is also L-true whereas (34) is not. Thus, we can re-write (34) and (35) as

(36) Scott is not L- equivalent to the author of Waverly

(37) Scott is L- equivalent to Scott.

If we compare (36) and (34) we find that although the entities which correspond to the individual expressions "the author of Waverly" and "Scott" are equivalent they are not L- equivalent. Furthermore, since a designator may, in general, be replaced by another one only if they are not merely equivalent but L- equivalent and since, by Leibniz's Law, only designators which denote the same entity are substitutable salva veritate, it follows that the entities which correspond to the individual expressions "Scott" and "the author of Waverly" are not one and the same entity. However, the extensions of "Scott" and "the author of Waverly" are the same, i.e., the individual Scott. In order to see that the extensions of "Scott" and "the author of Waverly" cannot be Carnap's individual entities it suffices to see that from (34) we can get

(38) $(\exists x)(x \text{ is equivalent to Scott})$

If the variable in (38) takes the individual Scott as its value then,

since there is only one individual Scott, there can only be one value for the variable in (38). But from (34) - (37) we get

(39) $(\exists x)(x \text{ is equivalent but not L- equivalent to Scott})$

(40) $(\exists x)(x \text{ is equivalent and L- equivalent to Scott}).$

Since one entity cannot satisfy the contradictory conditions (39) and (40), there must be at least two distinct entities and neither can be the individual Scott. The sentences (34) - (40) belong to the meta-language M^1 but since they are correct translations of sentences of S_2 what has been said about them can equally well be said of the sentences of which they are translations. The above argument is very similar to Quine's Morning Star - Evening Star argument in "The Problem of Interpreting Modal Logic".

The extensional object-language S_1 does not contain the translations of sentences (36) - (37) and (39) - (40). But it contains (34) - (35) and (38), or, rather, translations of these sentences. However, the meta-language M^1 will contain a true statement like

(41) 'S' L- designates Scott

(41) is a true statement in M^1 of both S_1 and S_2 . Since the relation of L- designation is such that an expression L- designates all entities L- equivalent to the entity to which the expression corresponds, from (41) we can get

(42) Scott is not L- equivalent to the author of Waverly.

which is a true statement in M^1 of S_1 . From (42) and (34) - (35) we can get (39) - (40). And since whatever holds of the designators and variables of M^1 also holds of the sentences of which they are translations we again see that the entities corresponding to individual

expressions of S_1 are not individuals but individual concepts. Moreover, as Davidson has pointed out we cannot say that the extensions of individual expressions are the entities designated by those expressions:

"It would be a mistake to imagine that the entities designated by expressions are the extensions of those expressions. Designation differs from L-designation, not in introducing new entities, but in introducing a new relation between expressions and the same entities. The extension of 'Scott' is intended to be that one individual, Scott; but 'Scott' designates each of the entities equivalent to the entity L-designated by 'Scott' (and of these there are at least two)."²⁹

We see, then, that although Carnap's method achieves a "reduction of entities" it is at the expense of extensions. As it has been pointed out by Quine this is a very radical move. Thus, he says

"The strangeness of the intensional language becomes more evident when we try to reformulate sentences such as these:

- (1) the number of planets is a power of three.
- (2) the wives of two of the directors are deaf.

In the familiar logic, (1) and (2) would be analyzed in part as follows:

- (3) $(\exists n)(n \text{ is a natural number. The number of planets} = 3^n)$
- (4) $(\exists x)(\exists y)(\exists z)(\exists w)[x \text{ is a director. } y \text{ is a director. } \neg(x = y). z \text{ is a wife of } x. w \text{ is wife of } y. z \text{ is deaf. } w \text{ is deaf}]$

But the formulation (3) depends on there being numbers (extensions, presumably classes of classes) as values of the bound variable; and the formulation (4) depends on there being persons (extensions, individuals) as values of the four bound variables. Failing such values, (3) and (4) would have to be reformulated in terms of number concepts and individual concepts. ... whose senses are more readily imagined than put into words."³⁰

Quine concludes by saying that although these considerations don't show that a language such as Carnap's is inadequate, they point to the unusual character of the language which comes out most clearly when dealing with examples such as (3) and (4) (above).

Thus far we have examined Frege's attempt to deal with the problems raised by Quine regarding quantification into opaque contexts. Apart from philosophical difficulties we found that Frege's solution to the problem of quantification into opaque contexts is unnecessarily complicated. It is natural to ask that since Frege's system leads to such complexities whether or not it might be modified in such a way as to restore some of its lost simplicity. Since Carnap's method of extension and intension may be regarded as a modification of Frege's system we were led to examine this method. We found that although Carnap's method can be used to deal with the puzzles raised by Quine regarding quantification into modal contexts, it does so only at the expense of extensions. We will now examine Smullyan's attempt to render quantification into modal contexts intelligible.

III

RECENT TREATMENTS OF THE PROBLEM

1. SMULLYAN AND SCOPE DISTINCTIONS: The distinguishing characteristics of Russell's theory of descriptions are the doctrines of descriptions as incomplete symbols, their contextual definition in use and their scope ambiguity. Let us concentrate on his discussion of scope ambiguity. According to Russell, the sentence "the present King of France is not bald" (or, for short " $\sim B(\exists x)(Kx)$ ") is ambiguous and needs to be disambiguated. The sentence is subject to two expansions according to whether the scope of the description is taken as large or small. That is to say, we may take the sentence as first of all saying that the

present king of France is bald and then denying it (small scope) or as simply affirming that the present king of France is not bald (has no hair). For the sentence $\sim B(\iota x)(kx)$ to be the contradictory of $B(\iota x)(kx)$, the scope of the description must be taken to be as small as possible. On this interpretation $\sim B(\iota x)(kx)$ abbreviates $\sim [(\exists c)\{(x)((kx) \equiv (x = c)) \& Bc\}]$. This is true whenever $(\iota x)(kx)$ does not have a descriptum. The other interpretation of $\sim B(\iota x)(kx)$ accords the description a large scope. Thus understood, $\sim B(\iota x)(kx)$ abbreviates $(\exists c)[(x)((kx) \equiv (x = c)) \& \sim Bc]$. This is false whenever $(\iota x)(kx)$ does not have a descriptum. The contradictory of $B(\iota x)(kx)$ is $\sim B(\iota x)(kx)$ with the description given small scope.

The case of negation may be extended to other truth-functions.

Thus, Russell says:

"The purpose of the following propositions is to show that, when $E! (\iota x)(\phi x)$, the scope of $(\iota x)(\phi x)$ does not matter to the truth-value of any proposition in which $(\iota x)(\phi x)$ occurs... The proposition can be proved generally when $(\iota x)(\phi x)$ occurs in the form $\chi(\iota x)(\phi x)$, and $\chi(\iota x)(\phi x)$ occurs in what we may call a 'truth-function', i.e., a function whose truth or falsehood depends only upon the truth or falsehood of its argument or arguments. This covers all the cases with which we are ever concerned. That is to say, if $\chi(\iota x)(\phi x)$ occurs in any of the ways which can be generated by the processes of *1 - *11, then, provided $E! (\iota x)(\phi x)$, the truth-value of $f\{[(\iota x)(\phi x)]\}$ is the same as that of $[\chi(\iota x)(\phi x)]$. $f\{\chi(\iota x)(\phi x)\}$."31

Russell's statement that "this covers all the cases with which we are ever concerned" should be taken to mean those cases with which we are concerned in truth-functional logic, i.e., extensional logic. When non truth-functional forms of composition occur, as with logical necessity and logical possibility then even proper descriptions (i.e., descriptions

with a descriptum) induce scope ambiguities. Smullyan was the first to make a note of this. He rightly argued that Quine's argument which tries to show that the principle of substitutivity of identity breaks down in modal contexts suffers from a fallacy of ambiguity - ambiguity of scope.³²

Let us re-write (1), (6) and (12), respectively, in the abstract form

(43) $\text{nec } (Fy)$

(44) $Y = (\exists x)(Gx)$

(45) $\therefore \text{nec } [F(\exists x)(Gx)]$

we see that (45) is ambiguous. It is to be disambiguated by placing a scope symbol somewhere. (45) could mean either

(46) $\text{nec } \{[(\exists x)(Gx)] \cdot [F(\exists x)(Gx)]\}$, i.e., $\text{nec } (\exists y)(x)[(Gx \equiv x = y) \wedge (Fy)]$

or

(47) $[(\exists x)(Gx)] \cdot \text{nec } [F(\exists x)(Gx)]$

Smullyan's point is simply that although (46) is false, it does not follow from (43) and (44). What does follow from (43) and (44) is the true sentence (47) which is the contraction of

(48) $(\exists y)(x)[(Gx \equiv x = y) \wedge \text{nec } (Fy)]$

Thus, Smullyan, in his reply to Quine, quite rightly observes that examples like (1), (6) and (12) which show the failure of substitutivity of identity in modal contexts, have to make use of descriptions and not just names. We can adjust matters by paying careful attention to distinctions of scope in Russell's theory of descriptions and also by slightly altering the theory to make it accommodate modal

logic. For Russell, as we saw from the quoted passage, the truth-value of a statement was not altered by changing the scope of the description occurring in it, unless the description in question failed to refer. In modal contexts, however, difference of scope can affect truth-value even where the description succeeds in naming. The following theorem does not hold in non truth-functional contexts

$$\text{EI } (\exists x)(\varphi x) \supset f \{[(\exists x)(\varphi x) \cdot \chi (\exists x)(\varphi x)]\} \equiv [(\exists x)(\varphi x) \cdot f \{ \chi (\exists x)(\varphi x) \}]$$

We have seen that (46) does not follow from (43) and (44). What does indeed follow by ordinary logic from (43) and (44) is (47), which is the contraction of (48). Moreover, there is no reason for not regarding (48) as true. Now, (48) is clearly a statement involving quantification into modal or opaque context. Smullyan assumes the intelligibility of statements like (48) without attempting to explain them. This is natural since Smullyan's article antedates the development of a satisfactory semantics for modal logic. Thus, in order to explain constructions like (48) we have to examine the semantics for quantified modal logic.

2. POSSIBLE-WORLDS SEMANTICS: As Kripke has shown we can use the notion of a possible world to interpret quantified modal logic.³³ We will start by giving a brief account of Kripke's semantics for quantified modal logic.

The systems of propositional modal logic which will be discussed are M (Von Wright), B ("Brouwershe system") and S_4 and S_5 (Lewis). In the language under consideration there are to be some unambiguously

constituted sentences. These sentences are the prime formulas or atoms and we denote them by capital Roman letters from late in the alphabet, such as "P", "Q", "R", ..., "P₁", "P₂" ... Distinct letters shall represent distinct atoms. Starting with the prime formulas or atoms and using the connectives $\&$, \sim and nec, we build other formulas A, B, C, ..., (called composite formulas or molecules), in the usual way. The prime formulas and the composite formulas together constitute the formulas of the language under consideration. The rules and axioms of the systems of propositional modal logic are as follows:

R0: If A is a theorem of the classical propositional calculus $\vdash A$

R1: If $\vdash A$ and $\vdash A \supset B$, then $\vdash B$

R2: If $\vdash A$ then $\vdash \text{nec } A$

A1: If $\vdash \text{nec } A$ then $\vdash A$

A2: $\vdash \text{nec } (A \supset B) \supset (\text{nec } A \supset \text{nec } B)$

A3: $\vdash A \supset \text{nec pos } A$

A4: $\vdash \text{nec } A \supset \text{nec nec } A$

The rules and axioms R0, R1, R2, A1, A2 belong to each system. We get the system B by adding A3; we get the system S₄ by adding A4 but excluding A3; and we get S₅ by adding both A3 and A4.

The semantics for propositional modal logic invokes the idea of possible worlds in which propositions are true or false. A model structure is a triple $\langle G, K, R \rangle$ where $G \in K$ and R a reflexive relation on K. We may think of G, says Kripke as the real world and of K as a class of possible worlds. R is to represent the relation of "relative possibility". A world H₂ is possible relative to H₁ if and only if every proposition true in H₂ is possible in H₁. We can get classes of valid

formulae belonging to different systems by varying the properties of R . For example, if R is transitive as well as reflexive we have as valid the theorems of Lewis's S_4 . A model on a model structure is just a function of two variables $\wp(A, H)$ that assigns T or F to atomic Formula A in world H. Truth-values for the complex formulas are defined in the following way

$V(A \& B, H) = T$ if and only if $V(A, H) = T$ and $V(B, H) = T$

$V(\neg A, H) = T$ if and only if $V(A, H) = F$

$V(\text{nec } A, H) = T$ if and only if $V(A, H^1) = T$ for every $H^1 \in K$ such that HRH^1

For quantified modal logic we add, for each positive integer n , a list of n -adic predicate letters (atomic formulas taken as 0-adic), together with individual variables and quantifiers. We define a quantificational model structure as a model structure $\langle G, K, R \rangle$ together with a function \mathcal{V} which assigns to each H , a set $\mathcal{V}(H)$, intuitively the set of individuals in H . A quantificational model on a quantificational model structure $\langle G, K, R \rangle$ is a binary function $\Phi(P^n, H)$, where the first variable ranges over n -adic predicate letters and H ranges over elements of K .

The truth-value of an atomic formula $P^n(x_1, \dots, x_n)$ in H relative to an assignment of objects a_1, \dots, a_n of $\bigcup_{H \in K} \mathcal{V}(H)$ to x_1, \dots, x_n can now be defined as follows:

$V(P^n(x_1, \dots, x_n), H) = T$ if and only if the n -tuple $\langle a_1, \dots, a_n \rangle \in \wp(P^n, H)$, otherwise $V(P^n(x_1, \dots, x_n), H) = F$, relative to the given

assignment. The assignments for complex formulas can be built up by induction. The induction steps for \sim , $\&$ and nec have already been

given. Consider the formula $A(x, y_1, \dots, y_n)$ where x and y_1 are the only free variables. Then $V(\Box A(x, y_1, \dots, y_n), H) = T$ relative to an assignment of b_1, \dots, b_n to y_1, \dots, y_n if and only if $V(A(x, y_1, \dots, y_n), H) = T$ for every assignment of a member a of $\mathcal{U}(H)$ to x . This completes the semantics for quantified modal logic.

We will illustrate the semantics by showing how the soundness of the various systems can be proved. Following Van Fraassen, we use \mathcal{T} as an index ranging over $\{m, b, 4, 5\}$, and use "System \mathcal{T} " to refer to system M if $\mathcal{T} = m$, B if $\mathcal{T} = b$, S_4 if $\mathcal{T} = 4$ and S_5 if $\mathcal{T} = 5$.³⁴ The languages under consideration are L_m, L_b, L_4 and L_5 . A valuation V is a mapping which assigns T or F to each atomic formula in each world. Since the syntax for each language is the same, we can define language $L_{\mathcal{T}}$ as the couple $\langle \text{Synt}, V_{\mathcal{T}} \rangle$, where $V_{\mathcal{T}}$ is the set of valuations of $L_{\mathcal{T}}$. We have to show that for each index \mathcal{T} , System \mathcal{T} is argument sound for $L_{\mathcal{T}}$.

Proof: We know from the classical propositional calculus that $R0$ and $R1$ are sound for $L_{\mathcal{T}}$.

Case (1). Suppose A is valid in the system \mathcal{T} , where $\mathcal{T} = m$. This means that every valuation V_m over the model structure $\langle G, K, R \rangle$ assigns T to A . That is to say, for every H^1 such that HRH^1 , $V(A, H^1) = T$. So by the valuation rule for "nec" we have $V(\text{nec } A, H) = T$. Hence $R2$ is sound.

Case (2). To show that $A1$ is valid in the system \mathcal{T} , where $\mathcal{T} = m, b, 4, 5$, we proceed as follows. We know that $V(\text{nec } A, H) = T$. This means that for all H^1 such that HRH^1 , $V(A, H^1) = T$. Moreover, since R is reflexive, we have as a special case $V(A, H) = T$.

Case (3). Similarly, we show that A2 is valid in the system \mathcal{T} , where $\mathcal{T} = a$. We know that $V(A \supset B, H^1) = T$ and $V(A, H^1) = T$ for every H^1 such that HRH^1 . Furthermore, we know by the valuation rule for " \supset " that in every world H^1 such that HRH^1 either $V(A, H^1) = F$ or $V(B, H^1) = T$. But $V(A, H^1) = T$. So $V(B, H^1) = T$. Moreover, since R is reflexive we have $V(B, H) = T$. Therefore, by the valuation rule for "nec" we have $V(\text{nec } B, H) = T$.

Case (4). To show that A3 is valid in the system \mathcal{T} , where $\mathcal{T} = b$, we suppose $V(A, H) = T$. Since R is symmetrical we know that for every H^1 , if HRH^1 , then H^1RH . So there is a possible world relative to H^1 in which A is true, namely, H. Therefore, $V(\text{pos } A, H^1) = T$. Since this is general and holds for all H^1 such that HRH^1 , we have $V(\text{nec pos } A, H) = T$.

Case (5). Consider the system \mathcal{T} , where $\mathcal{T} = 4$, and suppose $V(\text{nec } A, H) = T$. Because R is transitive we know that for all H^1 and H^{11} , such that H^1 is possible relative to H and H^{11} is possible relative to H^1 , $V(A, H^1) = V(A, H^{11}) = T$. Since this holds for all H^{11} such that H^1RH^{11} we have $V(\text{nec } A, H^1) = T$. Moreover, since it also holds for all H^1 such that HRH^1 , we have $V(\text{nec } \text{nec } A, H) = T$. Hence A4.

To illustrate the semantics for quantified modal logic we will show that

$$(49) \quad (\exists x) \text{ nec } Fx \supset (\exists x) \text{ nec } (Fx \vee Gx)$$

is a valid formula of Lewis' system S_5 .

Proof: To show that (49) is valid in the system \mathcal{T} , where $\mathcal{T} = 5$, we suppose $V((\exists x) \text{ nec } Fx, H) = T$. Assume for reductio ad absurdum that $V((\exists x) \text{ nec } (Fx \vee Gx), H) = F$. By the first hypothesis there is an

object a and for every H^1 such that HRH^1 , $V(Fx, H^1) = T$, relative to the assignment of a to x . But by the second hypothesis, there is an H^1 such that HRH^1 and $V((Fx \vee Gx), H^1) = F$, relative to the assignment of a to x . By the valuation rule for " \vee ", $V((Fx \vee Gx), H^1) = F$ if and only if $V(Fx, H^1) = F$ and $V(Gx, H^1) = F$ under an assignment of a to x . But by the first hypothesis $V(Fx, H^1) = T$ for every H^1 such that HRH^1 . We therefore have both $V(Fx, H^1) = T$ and $V(Fx, H^1) = F$ which is a contradiction. So (49) is valid.

We will now try to determine how constructions like

$$(48) (\exists y) (x) [(Gx \equiv x = y) \& \text{nec. } (Fy)]$$

can be explained by means of the semantics for quantified modal logic.

Before doing that, however, we have to make a distinction between abstract formal semantics for quantified modal logic and applied formal semantics for modal logic.

There are two ways of construing the semantical analysis of modal logic that Kripke has given us. Strictly speaking, what Kripke is offering us is a system of abstract formal semantics for modal logic. A model structure, for example, has no obvious connection with our ordinary modal notions. In Kripke's semantics a model structure is simply any ordered triple $\langle G, K, R \rangle$, where K is a set, $G \in K$ and R is a relation on K . K could be construed as a set of baseball players with G as the short-stop and R the relation "to the left of". Similarly, a quantified model structure is simply an ordered pair whose first member is a model structure, and the second member is a binary function $\psi(H)$ which assigns to each member H^1 of K a set of individuals—a set of baseball players, for example. Kripke's pure semantics does not, therefore, commit us to any philosophical

doctrine. The pure or abstract formal semantics does not give us a meaning for "nec" or tell us when a proposition is necessarily true or what it is for an object to have a property essentially. It only defines the notion of validity for each of the systems under consideration. For example, it tells us that (49) is valid in the system S_5 if and only if $((\exists x) \text{ nec } Fx \supset (\exists x) \text{ nec } (Fx \vee Gx), H) = T$ in every S_5 quantificational model on a quantificational model structure. So the abstract formal semantics only tells us what it is for a formula to be valid in a given system. It does not tell us what "nec" means. Nor does it tell us what a sentence like (48) means. For that we have to turn to applied formal semantics.

In the applied formal semantics a model structure $\langle G, K, R \rangle$ cannot be construed as just any triple where G is a member of K and R a relation on it. In applied formal semantics K will be a set of possible worlds where G is the actual world. And $\mathcal{V}(H)$ will assign that domain of objects to H which exist in H . Thus, for example, Kripke says

"Intuitively, we look at matters thus: K is the set of all 'possible worlds'; G is the 'real world'. If H_1 and H_2 are two worlds $H_1 R H_2$ means intuitively that H_2 is 'possible relative to' H_1 ; i.e., that every proposition true in H_2 is possible in H_1 ... We may impose additional requirements, corresponding to various 'reduction axioms' of modal logic: If R is transitive, we call $\langle G, K, R \rangle$ an S_4 -m.s.; if R is symmetric, $\langle G, K, R \rangle$ is a Bronwershe m.s.; and if R is an equivalence relation, we call $\langle G, K, R \rangle$ an S_5 -m.s."³⁵

One important difference, then, between abstract formal semantics and applied formal semantics is that in applied formal semantics more conditions are imposed on the notion of modelhood. For example, we have

already seen that $\mathcal{V}(H)$ assigns that domain of objects to H which exist in H . Furthermore, in applied formal semantics the members of the union of the domains of the possible worlds, i.e., $\bigcup_{H \in K} \mathcal{V}(H)$, are to have properties in the various worlds such that various predicates are to be regarded as being true of them, with respect to these worlds. And the same individual may have different properties in the different worlds. Thus, in applied formal semantics we get a meaning for "nec". For example a sentence like

(50) $(\exists x) \text{ nec } (x \text{ is rational})$

tells us that there exists an object in the actual world which has the property of being rational in every possible world. Hence, applied formal semantics gives us a meaning for statements involving quantification into modal or opaque contexts, such as (48).

Let us, however, examine the notion of an object having the same property in different possible worlds. To simplify things, we will suppose that there are just two objects a and b and just two possible worlds, one of which is our world, i.e., the actual world. Let us suppose, furthermore, that there are just two predicates 'F' and 'G' which are to be regarded as true or false of the objects a and b in the two possible worlds in the following way

Possible non-actual
world

Fa & Ga

Fb & Gb

Actual world

Fa & Ga

Fb & Gb

Since, in applied formal semantics, necessity is characterized in terms of truth in all possible worlds and possibility in terms of truth in some possible world, we have $\text{nec } Fa$, $\text{pos } Gb$, $\text{nec } Fb$ —, etc. And we

also have $(\exists x) \text{ nec } Fx$, $(\exists x) \text{ nec } \sim Fx$... etc. Thus "nec Fa" is to mean that "F" is true of a in all worlds which are possible with respect to the actual one (and since there is only one non-actual world it is true of a in that world), and " $(\exists x) \text{ nec } Fx$ " is to mean that there is some object x of which "F" is true in all these worlds. To make sense of this, we need some theory of individuation which accounts for identification of objects from one possible world to another. If, for example, one wants to keep track of an object from one possible world to another by the help of its attributes one might hold that every object has some attributes which it keeps in all worlds which are possible with respect to the actual one and which no other object has in any of these possible worlds. The problem of making such identifications (which has been called the problem of trans-world identity) and the difficulties which are raised by this problem will concern us in the next chapter. The difficulty faced by any theory of individuation which assumes the notion of one and the same individual existing in two worlds, can briefly be spelt out in the following way.

Let us suppose that Aristotle exists in some world H which is distinct from this world. Let us suppose, further, that in the world H, Aristotle is not the tutor of Alexander the Great. In the world H he may also lack other attributes or properties which he has in this world. But then how can we possibly identify Aristotle in the world H? It cannot be argued that we identify Aristotle in world H by means of the properties which he possesses in this world. For, it could well turn out that in the world H someone else possesses the properties which in fact characterize Aristotle in this world. But if we cannot

identify Aristotle in the world H, then we do not really understand the assertion that he exists there. And it seems to follow that trans-world identification doesn't make much sense at all and we will be led to something like Lewis' Counterpart theory according to which no object exists in more than one world.

In the next chapter we will argue that the problem of trans-world identity can be resolved by means of the notion of rigid designation. But, tentatively, we may conclude on the basis of the above argument that quantification into modal contexts doesn't make sense unless one can make sense of trans-world identification.

SUMMARY: In this chapter we have examined Quine's attack on modal logic by considering two proposed solutions for the difficulties arising out of quantification into modal contexts. Frege's theory of indirect speech and Carnap's modification of it were examined and found inadequate for handling the difficulties raised by Quine. Smullyan's solution for handling failures of substitutivity in opaque contexts by making scope distinctions was seen to be adequate for handling the difficulties raised by Quine in their original form. This followed from the fact that all the difficulties raised by Quine had revolved around the failure of the law of substitutivity of identity in modal contexts which rested on the fallacy of ambiguity exposed by Smullyan. However, the difficulties become apparent once again when we try to interpret quantified modal logic with the help of possible world semantics. In particular, the interpretation (i.e., the applied formal semantics) of quantified modal logic tends to give rise to the problem of trans-world identification which once again threatens the intelligibility of quantification into opaque contexts.

FOOTNOTES

¹The breakdown of the principle of substitutivity of identity in modal contexts had already been noted by Quine in "Whitehead and the rise of modern logic" in Schilpp, The Philosophy of Alfred North Whitehead (1941), pp. 141-142 and p. 148. In the same year, in his review of Russell's Inquiry into Meaning and Truth (Journal of Symbolic Logic, 6 (1941), pp. 29-30), Quine pointed out the difficulties of quantifying into belief contexts.

²Quine, W.V.O. "Three Grades of modal involvement" in The Ways of Paradox and other Essays, (Harvard University Press: 1976), p. 159.

³Ibid., p. 160.

⁴Ibid., p. 170.

⁵Quine, W.V.O. "Notes on existence and necessity" in Linsky (ed.) Semantics and the Philosophy of Language (University of Illinois Press: 1972), pp. 77-95.

⁶Ibid., p. 88.

⁷Church, A. Review of Quine's "Notes on existence and necessity", Journal of Symbolic Logic 8 (1948) pp. 45-7.

⁸Quine, W.V.O. "The problem of interpreting modal logic" Journal of Symbolic Logic, 12 (1947) pp. 43-48.

⁹Quine, W.V.O. "Reference and Modality" in Linsky (ed.) Reference and Modality (Oxford University Press, 1971), pp. 17-34.

¹⁰Ibid., p. 25.

¹¹Ibid., p. 26.

¹²Quine, W.V.O. Word and Object, (Technology Press: 1960), pp. 197-198.

¹³Quine, W.V.O. "Reply to Professor Marcus" in Ways of Paradox and other essays, op. cit., pp. 177-184.

¹⁴Ibid., p. 184.

¹⁵Quine, W.V.O. "Three Grades of modal involvement" in Ways of Paradox and other essays, op. cit., pp. 175-176.

¹⁶Quine, W.V.O. "Reference and modality" in Linsky (ed.) Reference and Modality, op. cit., p. 31.

¹⁷Frege, G. "On Sense and Reference" in Philosophical Writings of Gottlob Frege, translated and edited by P.T. Geach and M. Black (Oxford: Basil Blackwell, 1970), pp. 56-78. For Frege, a definite description would be regarded as a proper name. Thus, he says "... by 'sign' and 'name' I have here understood any designation representing a proper name, which thus has as its reference a definite object (this word taken in the widest range), but not a concept or a relation, ... the designation of a single object can also consist of several words or other signs. For brevity, let every such designation be called a proper name", Ibid., p. 57.

¹⁸It is not clear what Frege's reasons are for saying that the extension of a sentence is its truth-value. However, an argument for regarding the extension of a sentence to be a truth-value can be extracted from Carnap. Carnap in his Meaning and Necessity draws an analogy between truth-values of sentences and extensions of predicators. Thus, he says "A predicator of degree n is characterized by the fact that we must attach to it n argument expressions to form a sentence. Therefore, a sentence might be regarded as a predicator of degree zero. Let A_m and A_n be any predicators of degree n ($n \geq 1$); then ... A_m and A_n have the same extension if and only if $(x_1)(x_2) \dots (x_n)[A_m x_1 x_2 \dots x_n \equiv A_n x_1 x_2 \dots x_n]$ is true. If we stipulate that this, which applies originally only to $n \geq 1$, is to be applied analogously to sentences as predicators of degree zero, we find that two sentences A_m and A_n have the same extension if and only if $A_m \equiv A_n$ is true, hence if and only if A_m and A_n are equivalent." Meaning and Necessity (Chicago: University of Chicago Press, 1947), sec. 5, p. 26. Given that two designators have the same extension if and only if they are equivalent it is natural, in the case of sentences, to take the truth-values as extensions.

¹⁹Frege, G. "On Sense and Reference" in Philosophical Writings of Gottlob Frege, op. cit., p. 67.

²⁰For the sake of simplicity we will regard "wanted to know whether ..." as a single opacity inducing operator.

²¹Kaplan, D. "Quantifying In" in Reference and Modality, op. cit., p. 119.

²²Carnap, R. Meaning and Necessity, op. cit., section 30, p. 130.

²³Ibid., p. 131.

²⁴The sign " \equiv " takes the place of the identity sign between individual expressions. Between two n -place predicate expressions it is defined as standing for the same expression with variables inserted in the predicate places and the whole formula closed with universal

quantifier. Thus, for example, if A_m and A_n are two predicates of degree n ($n \geq 1$), then A_m and A_n are equivalent if and only if $(x_1) (x_2) \dots (x_n) [A_m x_1 x_2 \dots x_n \equiv A_n x_1 x_2 \dots x_n]$ is true.

²⁵ Carnap, R. Meaning and Necessity, op. cit., sections 1, 34, 41.

²⁶ Ibid., section 37.

²⁷ Davidson, D. "The Method of Extension and Intension" in Schilpp (ed.), The Philosophy of Rudolf Carnap (London: Cambridge University Press), pp. 311-349.

²⁸ Quine, W.V.O., as quoted by Carnap in Meaning and Necessity, op. cit., pp. 196-197.

²⁹ Davidson, D. "The Method of Extension and Intension" in The Philosophy of Rudolf Carnap, op. cit., p. 323.

³⁰ Quine, W.V.O. as quoted by Carnap in Meaning and Necessity, op. cit., p. 197.

³¹ Russell, B. & Whitehead, A. Principia Mathematica (Cambridge University Press: London, 1973), second edition, p. 184.

³² Smullyan, A. "Modality and Descriptions" in Reference and Modality, op. cit., pp. 35-40 and also his review of Quine's "The Problem of Interpreting modal logic" in Journal of Symbolic Logic, vol. 12 (1947), pp. 139-141.

³³ Kripke, S. "Semantical considerations on modal logic" in Reference and Modality, op. cit., pp. 63-73.

³⁴ van Fraassen, B. Formal Semantics and Logic. New York, 1971, pp. 144-151.

³⁵ Kripke, S. "Semantical considerations on modal logic" in Reference and Modality, op. cit., p. 64.

CHAPTER II

RIGID DESIGNATION

INTRODUCTION: In the last chapter we saw that to make sense of modal sentences like "nec Fa" and " $(\exists x) \text{ nec Fx}$ " we need some theory of individuation which will enable us to identify objects across possible worlds. This followed from the fact that we took "nec Fa" to mean that F is true of a in all worlds which are possible with respect to the actual world. There are, basically, two ways of dealing with this problem of trans-world identity which give rise to radically different solutions to this problem. One way of dealing with cross-world identity puzzles is by adopting what I shall call the Frege-Carnap-Kaplan method. This method of dealing with the problem of trans-world identity seems to me to be mistaken insofar as it leads us to posit counterparts of actual objects in other possible worlds. We will examine this method in the first section. It will be seen that counterpart theory is very closely allied with another theory which treats proper names as abbreviated descriptions. This is the view of the classical logical tradition as represented by Frege and Russell that a proper name simply was a definite description. Accordingly, we will examine this view in section II. The outcome of this investigation will be that proper names are not abbreviated or disguised descriptions. In the third section the theory which treats proper names as rigid

designators rather than as abbreviated descriptions will be presented and defended. We will argue that the notion of a rigid designator can be used to solve the problem of trans-world identity. The method of solving the problem of trans-world identity by treating proper names as rigid designators is due primarily to Kripke, and is the second method mentioned above. Finally, in the last section an important distinction will be drawn (which is again due to Kripke) between fixing the reference of a name and giving a synonym for it.¹

I

THE PROBLEM OF TRANS-WORLD IDENTITY: FREGE-CARNAP-KAPLAN METHOD

In the last chapter we saw that, according to Carnap, each designator has both an intension and an extension. Sentences have truth-values as extensions and propositions as intensions, predicates have classes as extensions and properties as intensions, terms have individuals as extensions and individual concepts as intensions, and so on. Moreover, two sentences are said to have the same intension if they are L-equivalent. Carnap's method of extension and intension was meant to be a modification of Frege's doctrine of indirect denotation. But Frege did not say what intensions were. In this respect Carnap did not differ from Frege. Both Carnap and Frege were concerned, rather, with the problem of trying to explain when two designators have the same intension.² Kaplan, on the other hand, proposes to understand propositions, i.e., the category of intensions appropriate to sentences as characteristic functions of sets of

possible worlds.³ These characteristic functions assign to each possible world one of the two truth values. This can be easily generalized. Thus, the intension of an individual expression will be the function which assigns to each possible world an individual in the domain of that world; the intension of a predicate will be a function which assigns to each possible world a set of individuals, and so on. In order to handle expressions which do not have an extension in a world we can divide the world into two parts: the entities which exist in that world and those which don't. This allows us to continue to treat intensions as functions from possible worlds to extensions without the consequence that such a function must always assign to each possible world an entity which exists in that world.

The problem which concerns us arises when we try to determine the intension of an open sentence. Consider, for example, the open sentence "x is the teacher of Alexander". We want to determine the intension of "x is the teacher of Alexander" under an assignment of values to the free variable. The function associated with the open sentence "x is the teacher of Alexander" is going to determine a proposition for every assignment of values to the variable. And different assignments of values to the variables might yield different propositions. Let us follow Kaplan and call the functions from individuals to propositions, propositional functions.⁴ Thus, Kaplan:

"The more subtle among you will have noted that this coining of the phrase 'propositional function' does nothing to clarify the notion. Let me review: we are clear on the proposition expressed by the closed sentence \mathcal{T} .. it is that function which assigns to each world w , the truth value of \mathcal{T} in w , and we are also clear on the simpler notion of a proposition .. any

function from worlds to truth values. We are clear on the simple notion of a propositional function ... a function from individuals to propositions (and we know what propositions are). But we are unclear on the notion the propositional function expressed by the open formula $\bar{\phi}$.⁵

Given that the intension of an expression is its extension in all possible worlds, in order to determine the propositional function associated with the open sentence "x is the teacher of Alexander", we have to determine the extension of this sentence under each assignment of a value to the variable "x". Let us try this in the case where the value of "x" is Aristotle. In the case of the actual world we can say that "x is the teacher of Alexander" is true of Aristotle. What about the other possible worlds? A possible world pretty much like the actual world except that in it Humphrey and not Nixon had won the election in 1968, would be a world in which the open sentence "x is the teacher of Alexander" would be true of Aristotle. What about a possible world in which Aristotle was not the pupil of Plato. How are we to decide in this case whether "x is the teacher of Alexander" is true of Aristotle? According to Kaplan we are to peep into this other world through his Jules Verne-o-scope .. "examine each individual, check his fingerprints, etc."⁶ and try to determine whether Aristotle is the teacher of Alexander in that world.

According to Kaplan, then, we are to try and identify individuals in different possible worlds in terms of qualities:

"This position might be associated with the bundle-- of qualities metaphysics insofar as it is associated with any particular metaphysical view. We so-to-speak look only at the clothes, and we identify an individual in terms of their strikingly similar manner of dress, i.e. their sharing of a large number

of prominent qualities. Our view of individuals in different worlds is through the Jules Verne-o-scope, which, you recall, enables us to compare fingerprints, ASCAP registeries, ... but does not allow us to see into any underlying bare particular ..."⁷

Now, let us go back and re-consider the question whether the open sentence "x is the teacher of Alexander" would hold of Aristotle in a world in which Aristotle was not the pupil of Plato. Let us call this possible world w_1 . We have to decide whether Aristotle in w_1 is the teacher of Alexander. Given that we have looked at the transcriptions from the Verne-o-scope we should, according to Kaplan, have no difficulty in deciding which individual in the world w_1 is to be connected with our Aristotle. It is of some importance to note that we can only decide which individual is connected with our Aristotle in w_1 and not which individual is identical with our Aristotle in w_1 . We cannot say that the Aristotle of w_1 is identical with our Aristotle because the Aristotle of w_1 is not the pupil of Plato whereas our Aristotle is. "And whatever you may think about the identity of indiscernibles, no sensible person would deny the indiscernibility of identicals"⁸ Kaplan calls the Aristotle of w_1 the counterpart of our Aristotle.⁹

According to Kaplan, then, the Aristotle of w_1 is the counterpart of our Aristotle because the difference between our Aristotle and the Aristotle of w_1 are regarded by us as minimal, whereas the similarities between the two is great. Another way to put it would be to say that since the difference between our world and the world w_1 is inessential we can regard the Aristotle of w_1 as a counterpart of our

Aristotle. However, as the disparity between our world and the other worlds grows it becomes more and more difficult to regard individuals in these other possible worlds as counterparts of our individuals, i.e., as counterparts of the individuals of our world. We have to make finer and finer discriminations between what is essential to our Aristotle and what is only accidentally true of him. In other words, we have to search for Aristotle's essence, that which will identify an individual of any possible world as the counterpart of our Aristotle. Now, one way we might do this is by searching for a description which speakers associate with their use of the name "Aristotle" and which expresses the essence of Aristotle.¹⁰ That there must be some description associated with a proper name is the view of the classical tradition represented by Frege and Russell. We will examine this view in the next section.

II

FREGE-RUSSELL VIEW

1. THE DESCRIPTION THEORY: In The Philosophy of Logical Atomism, Russell says that the only kind of word which can stand for a particular is a proper name.¹¹ Proper names are for Russell, by definition, words which stand for particulars. At the same time he maintains that what we ordinarily call names, like "Socrates" and "Aristotle" are really not names but abbreviations for descriptions. Thus, he says:

"The names that we commonly use, like 'Socrates', are really abbreviations for descriptions ... A name, in the narrow logical sense of a world whose meaning is a particular, can only be applied to a

particular with which the speaker is acquainted, because you cannot name anything you are not acquainted with ... we are not acquainted with Socrates, and therefore cannot name him. When we use the word 'Socrates', we are really using a description. Our thought may be rendered by some such phrase as, 'The master of Plato', or 'The philosopher who drank the hemlock', ..., but we certainly do not use the name as a name in the proper sense of the word."¹²

There are two important points to note about Russell's view. According to Russell names (in the proper sense) don't abbreviate descriptions. The sense of a name is the object or particular named. It follows from this view that the things which we normally call 'names' such as 'Socrates' and 'Aristotle' because they abbreviate descriptions are not really names in the proper sense. The second point to note is that names as they are ordinarily called do have a sense which is given by the description associated with the name. It is not clear, however, whether Russell would agree in thinking that the sense of a proper name, i.e., an ordinary proper name, is given by the description associated with the name. For one thing, Russell held that descriptions did not have sense. In 'On Denoting' he seems to hold the view that the sense-reference distinction was incoherent.¹³ When we talk about Russell's view it should be kept in mind that we differ from him in two respects. First, when we talk about names, we are to be taken to be talking about names in the ordinary sense. Secondly, we regard the sense of these names to be given by the description which they abbreviate.

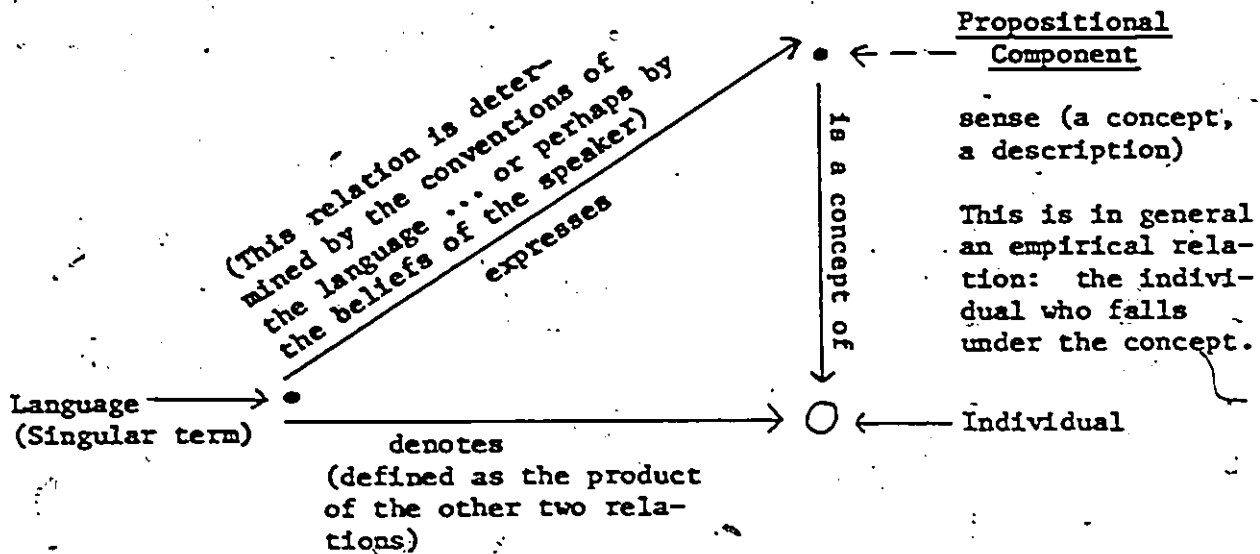
According to Frege, on the other hand, an ordinary proper name always has a sense, although it may sometimes lack a reference. A certain looseness concerning the sense of a proper name was recognized by

Frege. Thus, he says:

"In the case of an actual proper name such as 'Aristotle' opinions as to the sense may differ. It might, for instance, be taken to be the following: the pupil of Plato and teacher of Alexander the Great. Anybody who does this will attach another sense to the sentence 'Aristotle was born in Stagira' than will a man who takes as the sense of the name: the teacher of Alexander the Great who was born in Stagira. So long as the reference remains the same, such variations of sense may be tolerated, although they are to be avoided in the theoretical structure of a demonstrative science and ought not to occur in a perfect language."¹⁴

Frege seems to be saying two things. First of all, he is claiming that the sense of a proper name (the individual concept in Carnap's terminology) is given by a description so that the propositions expressed by the sentence 'Aristotle was born in Stagira' are different for two people who attach different senses to 'Aristotle'. Secondly, this difference is within tolerable limits because the propositions are not only about Aristotle, they also have the same truth-value so long as the different senses associated with 'Aristotle' present the same referent. So it is the sense of a proper name which determines its referent. This is what is meant by Frege's dictum that there is no road back from denotation to sense. Following Kaplan we can represent the situation schematically.¹⁵

Fregean Picture



This Fregean picture is meant to account for more than just singular terms, but since we are here concerned only with proper names, it will suffice for our purposes. What is important for our purposes is the fact that it is a description in a qualitative language which is the sense of a singular term and which determines the referent of the singular term. We will now examine the Russell-Frege view.

2. ARGUMENTS AGAINST THE FREGE-RUSSELL VIEW: The difficulty with the Frege-Russell view is that according to their theory even though different persons may attach different senses to the same name, any one person always attaches a unique sense to each name on all occasions, or even on any one occasion of its use. The reason for regarding this as a difficulty is that it is just bad psychology. When I use the name "Aristotle", for example, I do not have any one description in mind which determines the sense of the name "Aristotle" for me. But that is not all. Not only do I not have any one description in

mind when I use the name, I could not subsequently produce any one description which would give the sense which I attach to my use of the name "Aristotle". This remains true even if we grant that there is no description before my mind when I use the name "Aristotle". In some cases the Russell-Frege theory seems to be true but in most cases it fails.

In the case of some famous historical figures like Homer or Kerensky something like Russell's theory seems true. All I know about Homer is that he is the author of the Iliad. Regarding Kerensky, the description "the leader of the Russian Provisional Government" exhausts my information about him. However, if we turn to our friends and acquaintances we will find it very difficult, if not impossible, to give any one description which gives the sense of each of the names of our friends or acquaintances. Perhaps, some such considerations led Russell to distinguish between those whom we know "by acquaintance" and those whom we know only "by description". Thus, he says:

"A name in the narrow logical sense of a word whose meaning is a particular can only be applied to a particular with which the speaker is acquainted, because you cannot name anything you are not acquainted with. You remember, when Adam named the beasts, they came before him one by one, and he became acquainted with them and named them."16

However, it is quite clear that among some of the objects which I know "by description" are people like Aristotle and Plato who are similar in the relevant respect to some of the people I know by acquaintance. There is no unique definite description which I would be willing to substitute for the names "Aristotle" and "Plato" on all occasions of their use. The difference is not between those people whom we know 'by acquaintance' and those about whom we know 'by description' but between those about whom we

have a great deal of information and those about whom we have very little.

As mentioned above it is only in the case of a name like

"Homer" that Russell's theory seems to work. But even here it is not completely free from difficulties. There are some scholars who agree that there was a poet named "Homer" who wrote the Iliad and the Odyssey, while other scholars attribute the Odyssey to a poet who lived slightly after the author of the Iliad. If it is possible that the Odyssey and the Iliad were written by different people, then there is the problem whether Homer was the author of Iliad or the author of the Odyssey. Is it possible that Homer was neither the author of the Odyssey nor the author of Iliad but that he was a real person nonetheless? In this respect, then, the name "Homer" seems to function as Frege's theory holds names to function. But, of course, Frege and Russell's account of how the majority of ordinary proper names function, appears to be wrong. But if proper names are not abbreviations of some unique definite description, i.e., if proper names are not abbreviated descriptions then we have to ask ourselves what they are. To answer this question we have to examine the view of Saul Kripke that names are rigid designators. We will address ourselves to this task in the next section.

III

SEMANTICS OF DIRECT REFERENCE

1. RIGID DESIGNATION: The term "rigid designator" was used by Kripke to characterize those expressions which designate the same object in all possible worlds.¹⁷ According to Kripke, all proper names are rigid designators and so are many common nouns such as "gold" and "tiger",

etc. Kaplan has pointed out that there is a confusion in the idea that a proper name would designate nothing if the object which bears the name were not to exist.¹⁸ Now, there are two definitions given by Kripke of a rigid designator. The first definition is the following:

"Let us call something a rigid designator if in any possible world it designates the same object, a non-rigid or accidental designator if that is not the case. Of course we don't require that the objects exist in all possible worlds. Certainly Nixon might not have existed if his parents had not gotten married, in the normal course of things. When we think of a property as essential to an object we usually mean that it is true of that object in any case where it would have existed. A rigid designator of a necessary existent can be called strongly rigid."¹⁹

According to this conception of a rigid designator, a rigid designator cannot, at a given world, designate something which does not exist in that world. The second definition of a rigid designator which appears to be inconsistent with the one given above, appears a few pages later.

"When I say that a designator is rigid, and designates the same thing in all possible worlds, I mean that, as used in our language, it stands for that thing, when we talk about counter-factual situations ... It is in this sense that I speak of it as having the same reference in all possible worlds. I also don't mean to imply that the thing exists in all possible worlds, just that the name refers rigidly to that thing. If you say 'suppose Hitler had never been born' then 'Hitler' refers, here, still rigidly, to something that would not exist in the counterfactual situation."²⁰

The situation appears complicated because of Kripke's distinction between a rigid designator and a strongly rigid designator. According to the first definition a strongly rigid designator designates the same object in all possible worlds because a strongly rigid designator designates an object which exists in all possible worlds. A rigid designator, it

then seems, designates the same object in any possible world in which the object exists. According to the second definition, however, a rigid designator designates the same object in all possible worlds, even though the object might not exist in some possible worlds. The second definition of a rigid designator seems to suggest that in some counterfactual situations the sentence 'Hitler does not exist' would have been true, because what the name 'Hitler' would have designated would not have existed. It is probably some such apparent inconsistency that seems to be bothering Kaplan. However, let us look a little more carefully at this matter of rigid designation. In another article "Identity and Necessity", Kripke writes:

"What do I mean by a 'rigid designator'? I mean a term that designates the same object in all possible worlds. To get rid of one confusion which certainly is not mine, I do not use 'might have designated a different object' to refer to the fact that language might have been used differently ... what I mean by saying that a description might have referred to something different, I mean that in our language as we use it in describing a counterfactual situation, there might have been a different object satisfying the descriptive conditions we give for reference. So, for example, we use the phrase 'the inventor of bifocals', when we are talking about another possible world or a counterfactual situation, to refer to whoever in that counterfactual situation would have invented bifocals, not to the person whom people in that counterfactual situation would have called 'the inventor of bifocals'. They might have spoken a different language which phonetically overlapped with English in which 'the inventor of bifocals' is used in some other way. I am not concerned with that question here."²¹

In short, Kripke seems to be saying that a designator is used rigidly if and only if, when using it to talk about another possible world or counterfactual situation, we use it to refer to the same object that

it refers to in the actual world. If, however, in using it to talk about another counterfactual situation, we use it to refer to whoever would have satisfied certain conditions in that situation, the designator is non-rigid. Thus, for example, if in another possible world George Washington, not Ben Franklin had invented bifocals, then, if "the inventor of bifocals" is being used rigidly, it still refers to Ben Franklin in that world, if used non-rigidly it refers to George Washington. It is, in a sense, misleading to say that "proper names or, rather, rigid designators designate the same object in every possible world". The reason is that this suggests that there is no possible world in which the proper name, "Socrates", for example, denotes some individual other than Socrates. This suggestion, as Kripke points out, is misleading. It is misleading because Socrates does not have essentially the property of being named "Socrates". His parents could have named him "Aristotle"; had they done so Socrates would not have been named "Socrates". Rather, Kripke's claim is to be understood in the following way. To talk about Socrates in other counterfactual situations, we use the name "Socrates"—even if, in these situations his name is "Aristotle" or he has no name at all. Thus Socrates could have been named "Aristotle"; if that had been the case then his name would have begun with "A". Similarly, when we say 'suppose Hitler had not been born', we use the name 'Hitler' to talk about a counterfactual situation in which Hitler does not exist, i.e., we use it to talk about a counterfactual situation in which the name refers to the same object that it refers to in the actual world.

It might be objected that though the name 'Hitler' might denote

the same individual with respect to any possible world in which he exists, it certainly cannot denote him with respect to a possible world in which he does not exist. Thus, when we say 'suppose Hitler had not been born', we couldn't possibly be using the name 'Hitler' to talk about a counterfactual situation in which Hitler does not exist, because 'Hitler' in that situation does not denote anyone. This, however, is a mistake. So what if Hitler does not exist in the counterfactual situation which we are talking about, or if the inhabitants of a world in which Hitler doesn't exist have no name for him. We have a name for him, namely, 'Hitler'. It is our terms and sentences whose denotation and truth-value are being assessed with respect to the counterfactual situation in question.

If we keep the above considerations in mind we can say that a rigid designator is a designator which designates the same object in all possible worlds, namely, the object which it designates in the actual world.

We have seen that, according to Kripke, all proper names are rigid designators. Kripke's arguments for the view that proper names are rigid designators will be examined in the next sub-section.

2. KRIPKE ON PROPER NAMES: In "Naming and Necessity", Kripke argues for the thesis that proper names are rigid designators. He does this by trying to show that the meaning of a proper name is not equivalent to the meaning of some definite description, or even to the meaning of a cluster of descriptions. We have already seen that the description theory is false; i.e., a proper name is not an abbreviated

description. One consequence of this view is that the reference of a name is not determined by some uniquely identifying mark, some unique property satisfied by the referent and known or believed to be true of that referent by the speaker. If, for example, the name "Aristotle" was a mere abbreviation for the description "the teacher of Alexander", then one could determine the reference of the name "Aristotle" by determining which object (if any) is unique in having the property of being a teacher of Alexander. But since the name "Aristotle" is not short for "the teacher of Alexander", the reference of the name is not determined by the uniquely inhering property given by the description.

One way of circumventing this difficulty is to say that although we cannot substitute a particular description for the name, we can substitute a cluster of descriptions for the name. In this case, then, the meaning of a name is not given by some particular description but by a cluster of descriptions. Thus, for example, Wittgenstein:

"Consider this example. If one says 'Moses did not exist', this may mean various things. It may mean: the Israelites did not have a single leader when they withdrew from Egypt—or: their leader was not called Moses—or: there cannot have been anyone who accomplished all that the Bible relates of Moses— ... But when I make a statement about Moses,—am I always ready to substitute some one of those descriptions for 'Moses'? I shall perhaps say: by 'Moses' I understand the man who did what the Bible relates of Moses, or at any rate, a good deal of it. But how much? Have I decided how much must be proved false for me to give up my proposition as false? Has the name 'Moses' got a fixed and unequivocal use for me in all possible cases?"²²

Among others who have held the cluster of descriptions view are N.L. Wilson and J. Searle.²³ However, for the present when we talk about

the cluster theory of names we are to be understood to be talking about the theory which holds that the meaning of a name is equivalent to the meaning of some cluster of descriptions.

Kripke's thesis that proper names are rigid designators consists in showing that the meaning of a proper name is not equivalent to the meaning of some particular description or even to the meaning of a cluster of descriptions. Since we have already disposed of the description theory, let us concentrate on Kripke's arguments against the cluster theory of names.

Consider, for example, the name "Aristotle". Let us assume for the moment that Aristotle did not do any of the things that he is claimed to have done, i.e., he did not study philosophy, and was not the teacher of Alexander, etc. Now, according to Kripke, a situation in which Aristotle did not do any of the things which he is in fact claimed to have done would be described as a situation in which Aristotle did not do any of those things. In describing the situation we think of one and the same person, namely Aristotle, and say of him that he might not have done any of those things. It is a part of the description of this counterfactual situation or possible world that in that world Aristotle did not do any of the things that he is in fact claimed to have done.

It is important to note that in thinking about a counterfactual situation and in describing what would have happened to Aristotle in this counterfactual situation we are not giving a purely qualitative description of Aristotle. What would constitute a purely qualitative description of Aristotle? Consider, for example, the name "Nixon".

Kripke says

"... if someone makes the demand that every possible world has to be described in a purely qualitative way, we can't say, 'suppose Nixon had lost the election', we could say instead 'suppose a man with a dog named Checkers, who looks like a certain David Fry impersonation, is in a certain possible world and loses the election.'"²⁴

A demand for a purely qualitative description of Aristotle or Nixon is a demand for a set of descriptions which exhaust the meaning of the names "Aristotle" and "Nixon". And this is closely connected with the demand for providing a purely qualitative description of a counterfactual situation. Thus, if we were to give a purely qualitative description of the situation in which Aristotle did not do the things which are commonly attributed to him, for example, did not teach Alexander, then we could not say 'Suppose Aristotle did not teach Alexander'. Rather, we would have to search for a cluster of descriptions which exhausts the meaning of the name "Aristotle". Once we had done that we could say something like "Suppose a man who was a pupil of Plato and one of the greatest philosophers of our time didn't teach Alexander", or something to that effect, where the descriptions "One of the greatest philosophers of our time" and "the man who was a pupil of Plato" are part of the meaning of the name "Aristotle".

Kripke criticizes the cluster theory of descriptions by criticizing the attempt to look for purely qualitative descriptions of individuals and counterfactual situations. Thus, in talking about Nixon he says:

"Why need we make this demand? That is not the way

we ordinarily think of counterfactual situations. We just say 'suppose this man had lost'. It is given that the possible world contains this man, and that in that world, he had lost. There may be a problem about what intuitions about possibility come to. But if we have such an intuition about the possibility of that (this man's electoral loss), then it is about the possibility of that. It need not be identified with the possibility of a man looking like such and such, or holding such and such political views, or otherwise qualitatively described, having lost. We can point to the man and ask what might have happened to him, had events been different."25

The general picture, then, against which Kripke is arguing is that it is a precondition of understanding the use of a proper name to refer to an object in a counterfactual situation that one is prepared to give some uniquely individuating property, or properties of the object. The description theory gives a uniquely individuating property by means of some particular description associated with the use of a proper name; and the cluster theory of names attempts to give a set of properties by means of a cluster of descriptions associated with the use of a name. Both theories give us qualitative necessary and sufficient conditions for using a name to refer to an object in a counterfactual situation. The acceptance of the description or the cluster theory of names leads naturally to the demand that every possible world has to be described in a purely qualitative way. For, each possible world contains a number of individuals with certain properties and certain relations to each other. We are supposed to use these properties and relations to decide which member (if any) of a given possible world is identical with a given member of another possible world. It is the attempt to try and identify individuals in

possible worlds by means of properties and relations, i.e., it is the attempt to provide purely qualitative criteria of identity for individuals that gives rise to the problem of trans-world identity. Thus, for example, let us suppose that Aristotle exists in some possible world (let us call it W) distinct from our world, a world in which, let us say, he did not teach Alexander. In the world W he may also lack other properties which he has in this world; for example, he may not have studied philosophy or been the pupil of Plato. How would we locate Aristotle in the world W? The description theorist would say that it cannot be done. If Aristotle in the world W does not have the unique property, i.e., the essence, which is given by means of some particular description associated with the name "Aristotle" in our world, then there is no way of telling which individual in the world W is our Aristotle. The cluster theorist would say that if some individual in W has a sufficient number of properties which are given by means of a cluster of descriptions associated with our use of the name "Aristotle" in this world, then that individual is Aristotle.

Kripke, on the other hand, wants to argue that it is our ability to refer (rigidly) to Aristotle that makes trans-world identifications unproblematic in these cases. Thus, he says:

"Those who have argued that to make sense of the notion of rigid designator, we must antecedently make sense of 'criteria of transworld identity' have precisely reversed the cart and the horse; it is because we can refer (rigidly) to Nixon and stipulate that we are speaking of what might have happened to him (under certain circumstances), that 'transworld identifications' are unproblematic in such cases."²⁶

As it stands Kripke's position is not wholly free from difficulties.

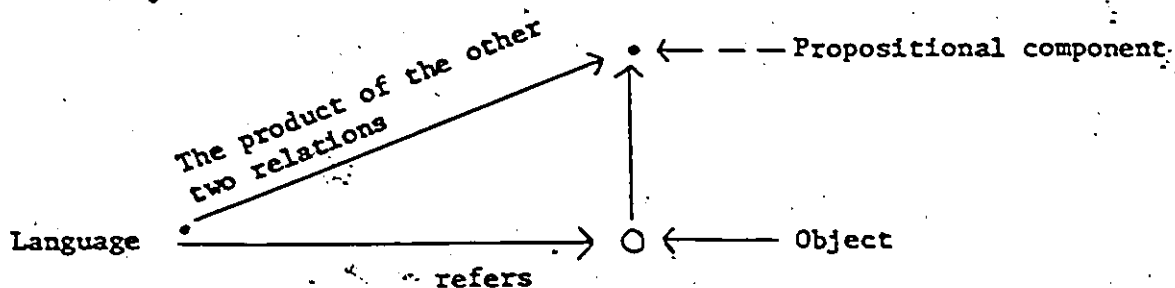
Kripke says that we can refer rigidly to Nixon and stipulate what might have happened to him. But surely, we cannot stipulate that Nixon might have been an inanimate object or that Nixon might have been a chimpanzee. If Nixon is in fact a human being then it seems intuitively bizarre to say that there might be counterfactual situations in which Nixon was not a human being or in which he was a chimpanzee. Let us therefore take it that it is a necessary fact about Nixon that he is a human being, that he is not an inanimate object, etc. But this difficulty can be easily handled. We have to slightly modify the definition of a rigid designator. Let us recall that our definition of a rigid designator was that it designates the same object in all possible worlds. To accommodate the essential properties of an object we have to modify this definition in the following way. A rigid designator is a designator which picks out the same object by essential traits in all possible worlds.²⁷ The advantages that accrue from this modification is that we don't have to say that the essential properties of Aristotle, for example, are given by some cluster of descriptions associated with the name "Aristotle". Instead we can say that the essential traits of Aristotle are captured by the rigid designator, namely, the name "Aristotle". But, even the above reformulation of the notion of a rigid designator will not do. The reason for this is that the above reformulation of the notion of a rigid designator does not, in any way, satisfy the requirement that there be some sufficient conditions for Aristotlehood. It merely satisfies the requirement that there be some necessary conditions for Aristotlehood which we can spell out. However, these necessary

conditions will not enable us to pick out an individual in a possible world. For example, that it is a necessary fact about Aristotle that he is a human being is of no use to us in picking out Aristotle or distinguishing him from other members of a possible world. If it is to be possible for us to stipulate that it is Aristotle and not someone else that we are talking about in a counterfactual situation, then this must be so because the proper name "Aristotle" expresses the individual essence of Aristotle. We will examine this notion in the next sub-section. But for the time being we can simply note that it is because proper names like "Aristotle" express the essence of Aristotle that we can stipulate that in talking about what would have happened to Aristotle in a counterfactual situation we are talking about what would have happened to him. Or to put it another way it is because a proper name like "Aristotle" expresses the essence of Aristotle that we can refer (rigidly) to Aristotle and stipulate that we are speaking of him in a certain counterfactual situation. Moreover, as it will become clear in the next sub-section, our ability to refer (rigidly) to Aristotle, for example, does not consist in our being able to spell out a set of qualitative necessary and sufficient conditions for Aristotlehood.

To sum up. The general picture which emerges from Kripke's treatment of proper names is one which is completely opposed to the Fregean picture of how the reference of proper names is secured. According to Frege, the reference of a name is determined via the sense of the name which is given by some description associated with the name. For Frege there is no road back from reference to sense.

According to Kripke, however, since the meaning of a name is not equivalent to the meaning of some description or a cluster of descriptions, the reference of a name is not determined by the sense of the name. The picture which emerges from the above discussion of proper names may be represented schematically as follows:²⁸

DIRECT REFERENCE PICTURE



3. PROPER NAMES AND ESSENCES: We have seen above that our definition of a rigid designator as a designator which picks out the same object by essential traits in all possible worlds, will not do. For essential properties can be of no use in providing a criterion of trans-world identification. In particular, essential properties which are common to several individuals will not serve to distinguish one individual from others. We, thus, had to fall back on the notion of an individual essence, i.e., that property which a particular individual alone has in every possible world. It does not seem to be too difficult to specify one such property in the case, for example, of Aristotle. One such property would be Aristotle's being identical with Aristotle. However, this property will not help us to pick out Aristotle in all possible worlds. Our problem is to identify Aristotle across possible

worlds and this cannot be done by finding the object which is identical with Aristotle in each possible world, for, our problem is how to do that. We thus see that not even some properties which Aristotle alone has in every possible world can be used to identify Aristotle in each possible world. We must, at this point, make a distinction between two issues which arise in connection with the problem of trans-world identification. It is commonly assumed that when we ask for a criterion of trans-world identification we are dealing with an epistemological issue. How can we identify the same object in different possible worlds given that the object has different properties in different worlds? If we embrace some form of essentialism, then, our answer will be that an individual is identified across possible worlds by means of its individual essence, i.e., we must look for the individual essence, where the individual essence is conceived as a set of qualities which we have to be able to recognize in order to identify the object in different worlds. But is it really the epistemological issue with which we are concerned when we are dealing with the applied formal semantics for quantified modal logic? I don't think so. What we have to do in order to understand the applied formal semantics for quantified modal logic is to make sense of the idea of one and the same object in different possible worlds. In other words, we have to provide an answer to the question "What does it mean to say that one individual in a world w_1 is the same as another in a different world w_2 ?" But this question is quite different from the question "How would we identify the same object in different possible worlds?" This latter question is an epistemological question and we

don't have to answer this question before we can say that we have an intuitive understanding of the semantics for quantified modal logic.

It is important to keep the above distinction in mind because it is the difficulty involved in providing an answer to the epistemological question which is usually cited to discredit quantified modal logic. On the other hand, the question "What does it mean for an individual in a possible world to be identical with an individual in another possible world?" which is really the only question which needs to be answered if quantified modal logic is to make sense, can very easily be answered. An object in a possible world is identical with another object in a different possible world if and only if the second object has all the essential properties of the first object. Included in the essential properties is the individual essence of the object. In other words, the question "What does it mean to say that an individual in a possible world is identical with another individual in a different possible world?", can be answered very easily and the answer to this question shows that the problem of trans-world identity does not pose any problems for the applied formal semantics for quantified modal logic.

The epistemological question, i.e., the question "How would we identify individuals across possible worlds?" is a question which we will not attempt to answer in this thesis. But it should be clear that the absence of an answer to this question does not mean that the applied formal semantics for quantified modal logic is unintelligible. For, we must make a distinction between the claim that we are unable to provide a satisfactory criterion which will enable us to answer

the question as to whether or not two individuals in different possible worlds are one and the same and the claim that such identifications are senseless in the absence of such a criterion. This latter claim does not follow from the former and it is false. Simply because we are unable to articulate a criterion for the re-identification of individuals across possible worlds does not mean that we cannot make sense of these re-identifications. As we have already seen we use proper names to make such re-identifications successfully. And what enables us to use proper names for the successful performance of this task is the fact that proper names are rigid designators. However, we have to slightly modify our definition of a rigid designator. We defined a rigid designator as a designator which picks out the same object in all possible worlds by means of its essential traits. But we have seen that the essential properties of an object are not sufficient to make sense of the idea of trans-world identity. Accordingly, we have to include the individual essence, i.e., a property or set of properties which an object alone possesses. If we take the necessary properties of an object to be its essential properties together with its individual essence, we can re-define the notion of a rigid designator along the following lines. A rigid designator is a designator which picks out the same object by means of its necessary properties in all possible worlds. We have seen that we don't have to spell out what these properties are in order to make sense of trans-world identifications; that we can make intuitive sense of trans-world identifications without being able to specify a criterion of trans-world identity. The above view is very similar to Kripke's

position regarding the problem of trans-world identity. Thus, for example Kripke says:

"... even if there were a purely qualitative set of necessary and sufficient conditions for being Nixon, the view I advocate would not demand that we find these conditions before we can ask whether Nixon might have won the election ... we can simply consider Nixon and ask what might have happened to him had various circumstances been different ... Notice this question ... is not epistemological ... It's a question about, even though such and such things are the case, what might have been the case otherwise."²⁹

What Kripke is saying is that we can use rigid designators to refer to Nixon, for example, in order to ask what might have happened to him under certain circumstances, without being able to spell out the necessary properties required for Nixonhood. And what makes this possible is the fact that we are not asking an epistemological question.

To sum up. We have seen that to make sense of the problem of trans-world identity we don't have to answer the epistemological question "How would we identify the same object in different possible worlds?" For, it cannot be maintained that in the absence of an answer to the epistemological question we cannot understand the applied formal semantics for quantified modal logic. All that is required for us to understand the applied formal semantics is that we can make sense of the idea of one and the same individual in different possible worlds, i.e., that we can provide an answer to what it means to say that one object in a possible world is the same as another in a different world. However, the answer to this question is very simple. An object in a possible world is the same as another in a different world if and only if the second object has all the necessary properties

of the first. Moreover, the necessary properties of objects are captured by rigid designators and it is because rigid designators express necessary properties that we can use them to refer to objects in counterfactual situations without spelling out the qualitative necessary and sufficient conditions for the identity of these objects.

It is of some importance to note that applied formal semantics, in the style of Kripke, is indeed committed to some form of essentialism. For, we have seen that to understand the semantics for quantified modal logic we have to make sense of the idea of one and the same individual in different counterfactual situations. Moreover, an object in a possible world is the same as another in a different possible world if and only if the second object has all the necessary properties of the first. There is a sense, then, in which modal logic is committed to essentialism. But we don't have to accept Quine's conclusion: "so much the worse for quantified modal logic". Quine rejects essentialism because he regards it as a nonsensical metaphysical doctrine. But, what, exactly are Quine's reasons for rejecting essentialism? In a passage which we have already quoted in the last chapter, Quine argues that a sentence such as

(1) $(\exists x)(x \text{ is necessarily greater than } 7)$

is meaningless. Thus, he says:

"Whatever is greater than 7 is a number, and any given number x greater than 7 can be uniquely determined by any of various conditions, some of which have ' $x > 7$ ' as a necessary consequence and some of which do not. One and the same number x is uniquely determined by the condition

(2) $x = \sqrt{x} + \sqrt{x} + \sqrt{x} + \sqrt{x}$

and by the condition

(3) There are exactly x planets

but (2) has ' $x > 7$ ' as a necessary consequence while (3) does not. Necessary greatness than 7 makes no sense as applied to a number x ; necessity attaches only to the connection between ' $x > 7$ ' and the particular method (2), as opposed to (3), of specifying x ."30

In the above passage Quine seems to be arguing that whether an object has a property essentially or necessarily depends on our method of specifying the object. Thus, for example, if the number 9 is uniquely determined by condition (2), then it will have the property of being necessarily greater than 7. But the number 9 could also be picked out as the number of planets. And in that case it will not have the property of being necessarily greater than 7. Hence, whether 9 has a certain property necessarily or contingently would depend on the way it is described.

It should now be clear where the flaw is in Quine's argument. If we refer to the number 9 by means of the description "the number of planets", then, of course, if things had been different the number of planets might not have been 9. So the term "the number of planets" does not designate the same object in all possible worlds, i.e., is not necessarily a rigid designator. On the other hand, the name "9" is just a name of the number 9. When we ask whether it is necessary or contingent that 9 is greater than 7, we are asking whether in some counterfactual situation 9 would in fact be less than or equal to 7. In other words, we are asking of the number 9 whether it has the property of being greater than 7 necessarily or contingently. Thus, if Quine thinks that it does not make any sense to talk about necessity attaching to the number 9, then he is wrong. It is because the term

"9" is a rigid designator and can be used to refer (rigidly) to the number 9 in all possible worlds, that we can ask of the number 9 whether it is necessarily greater than 7 or not. Note that we are not asking an epistemological question. We are not asking how we could come to know that the number 9 is necessarily greater than 7. We are simply considering the meaningfulness of the notion of a necessary property. And we are arguing that it makes sense to say that the number 9 has a property that can be held to be essential. To the extent that we understand statements like "9 is necessarily greater than 7", we can make sense of the idea of one and the same object in different possible worlds. And this understanding is all that is required to give intuitive sense to the applied formal semantics for quantified modal logic. All the better for quantified modal logic.

IV

ON FIXING THE REFERENCE vs. SUPPLYING A SYNONYM

We have seen that Kripke argues for the thesis that names are rigid designators by arguing that the meaning of a name is not equivalent to the meaning of some particular description or even a cluster of descriptions. Thus, for example, Kripke says:

"So suppose we say, 'Aristotle is the greatest man who studied under Plato'. If we used that as a definition, the name 'Aristotle' is to mean 'the greatest man who studied under Plato'. Then of course in some other possible world that man might not have studied under Plato and some other man would have been Aristotle. If, on the other hand, we merely use the description to fix the referent

then that man will be the referent of 'Aristotle' in all possible worlds."³¹

Kripke's argument here is meant to demonstrate the possibility of introducing a rigid designator by citing a definite description, not to give a definitional equivalent but only to "fix the reference". In the case of a proper name if it is specifically intended and stipulated that the introduction of the name is to be the introduction of a rigid designator by means of a description which fixes its referent, then there is no reason to suppose that the stipulation cannot accomplish what is intended. The situation we are concerned with is one in which a name is introduced by means of a description which fixes its referent and to which the name is pegged for some time. There is no reason to suppose that what we would conclude about this sort of situation should have implications for what we would say about, say, the present use of the names "Aristotle" or "Plato". We can maintain that these names are rigid designators and yet deny that they are now pegged to some definite description that fixes the referent. Consider, for example, a name such as "Neptune". This name was first introduced by means of a description which was used to fix its reference.

"An even better case of determining the reference of a name by description ... is the discovery of the planet Neptune. Neptune was hypothesized as the planet which caused such and such discrepancies in the orbits of certain other planets. If Leverrier indeed gave the name 'Neptune' to the planet before it was ever seen, then he fixed the reference of 'Neptune' by means of the description mentioned."³²

It is, however, certain that it is no longer tied to the description that was used to introduce it. This is shown by the fact that if we

were to ask an astronomer now how he knows that Neptune causes the perturbations in the orbit of Uranus, he would be treated to a discussion of such things as astronomical observations, details of orbits, etc. whereas just after Leverrier's stipulation the same question would have received some such reply as "That is just what we call 'Neptune'".

Now, Donnellan and Kripke have advanced theories of naming according to which someone who held that a name had originally been introduced as a rigid designator by means of one description, has no reason to say that now, when it has been disengaged from that description, there is some other description that fixes its referent. Instead, according to them, there is another relationship that can hold between the user(s) of a name, the name and the thing named in virtue of which the thing is named and that does not involve the thing satisfying any description associated with the name. Thus, Kripke speaks of a causal relationship;³³ Donnellan has talked about a relationship of being involved in an historical explanation of the use of a name.³⁴ A name, according to some version of such a theory, will be a rigid designator, but not because there is some definite description that fixes the reference of the name. Against Donnellan and Kripke we will argue that the breaking away of a name from the description which was used to introduce it is a case in which the introducing description has come to be merely one member of a corpus of assertions describing the individual (where the referent is determined by being that which satisfies some proportion of the descriptions). It is of some importance to note that we will not be arguing that the name is simply an

abbreviation for the corpus of utterances which determine the reference of the name. If that was the case then the name could not in general be regarded as a rigid designator. In order to preserve the rigidity of names we will argue that the breaking away of the name from the description which was used to introduce it is a case where the introducing description has come to be merely one member of a set of assertions which determine the reference of the name. However, prior to our presentation of a theory of names which preserves the rigidity of names we will examine Donnellan's and Kripke's theories of naming. This will be the task of the next two chapters.

SUMMARY: In this chapter we saw that the problem of trans-world identity can be solved by distinguishing between the question "How would we identify the same individual across possible worlds?" and the question "Does it make sense to talk about one and the same individual across possible worlds?" and also by treating proper names as rigid designators. To treat proper names as rigid designators is not to regard them as being equivalent in meaning to some particular description (the Frege-Russell view), or to some cluster of descriptions (the cluster theory of names). To say, for example, that there is a possible world in which Aristotle did not teach Alexander implies that, in that world, the reference of the name "Aristotle" is not determined in the way in which it is determined in the actual world; rather it is determined in that world via its reference in the actual world, i.e., as being to the same man who is its referent in the actual world. The reference of a name in the actual world is determined,

initially, by means of a description which fixes the reference of the name but is not synonymous with it. The problem to which we now have to address ourselves is how the reference of a name is determined once it has been disengaged from the description which was used to introduce it.

FOOTNOTES

¹According to M. Dummett the distinction between fixing the reference of a name by citing a definite description and giving a synonym for the name is due to Geach. See, for example, Dummett, M. Frege: Philosophy of Language (Duckworth & Co.: London, 1973), pp. 111-112.

²According to Kaplan, Carnap proposed that intensions appropriate to sentences be regarded as sets of possible worlds. However, it is difficult to find a passage in Carnap where he explicitly says this. Cf. Kaplan, D. "Trans-world Heir Lines". Paper presented at a joint symposium of A.P.A. in Chicago, 1967, p. 3.

³Kaplan, D. "Trans-world Heir Lines," op. cit., p. 4.

⁴Ibid., p. 5.

⁵Ibid., p. 6.

⁶Ibid., p. 7.

⁷Ibid., p. 14.

⁸Ibid., p. 15.

⁹The founder of contemporary counterpart theory is D. Lewis. He circumvents the problem of trans-world identity by insisting that nothing inhabits more than one world, thus substituting the counterpart relation for trans-world identity. Thus, he says "... something has for counterparts at a given world those things existing there that resemble it closely enough in important respects of intrinsic quality and extrinsic relations, and that resemble it no less closely than do other things existing there", Lewis, D. Counterfactuals (Basil Blackwell: Oxford, 1973), pp. 39-41. See also his "Counterpart Theory and Quantified Modal Logic", Journal of Philosophy 65 (1968) pp. 113-126.

¹⁰The view advocated by Kaplan where the essence of an object is not given by a fixed final description would come closer to the view advocated by Frege, below, rather than Russell.

¹¹Russell, B. Russell's Logical Atomism, ed. D. Pears (Collins Sons & Co.: London 1972), p. 55.

¹²Ibid., p. 56.

¹³Russell, B. "On Denoting". Reprinted in Logic and Knowledge, ed. R.C. Marsh (George Allen and Unwin: London, 1956), pp. 48-51.

¹⁴ Frege, G. "On Sense and Reference" in Translations from the Philosophical Writings of Gottlob Frege, ed. P. Geach and M. Black (Oxford: Basil Blackwell, 1952), p. 58, n.

¹⁵ Kaplan, D. (1977). Demonstratives (Draft #2) mimeographed; UCLA Philosophy Department, p. 4.

¹⁶ Russell, B. Russell's Logical Atomism, op. cit., p. 56.

¹⁷ Kripke, S. "Naming and Necessity" in Semantics of Natural Language, ed. D. Davidson and G. Harman (Dordrecht, 1972) pp. 253-355.

¹⁸ Kaplan, D. Demonstratives, op. cit., pp. 11-12.

¹⁹ Kripke, S. "Naming and Necessity", op. cit., pp. 269-270.

²⁰ Ibid., pp. 289-290.

²¹ Kripke, S. "Identity and Necessity" in Identity and Individuation, ed. M. Munitz (New York University Press: New York, 1971), p. 145.

²² Wittgenstein, L. Philosophical Investigations, translated by G.E.M. Anscombe (Basil Blackwell: Oxford, 1974), section 79.

²³ Wilson, N.L. "Substances without Substrata", The Review of Metaphysics, 12 (1959), pp. 521-539 and Searle, J. "Proper Names", Mind 27 (1958), pp. 166-173. Wilson's cluster theory of names, as we will see in the fourth chapter, is quite different from Wittgenstein's and Searle's theories.

²⁴ Kripke, S. "Naming and Necessity" in Semantics of Natural Language, op. cit., pp. 267-268.

²⁵ Ibid., p. 268.

²⁶ Ibid., p. 270.

²⁷ There are passages in Kripke's "Naming and Necessity" where Kripke seems to be aware of the fact that we cannot stipulate that there is a possible world in which Nixon, for example, is not a human being.

²⁸ This picture of direct reference can be found in Kaplan's manuscript on "Demonstratives", p. 4. We have, in order not to beg the issue, left unexplained the relation of reference.

²⁹ Kripke, S. "Naming and Necessity" in Semantics of Natural Language, op. cit., pp. 268-269.

³⁰Quine, W.V.O. "Reference and Modality" in Reference and Modality, ed. Linsky, L. (Oxford University Press: Oxford, 1971), p. 25.

³¹Kripke, S. "Naming and Necessity" in Semantics of Natural Language, op. cit., p. 276.

³²See also Kripke, S. "Naming & Necessity", op. cit., p. 345, n. 16 where Kripke says "... free variables can be used as rigid designators of unspecified objects".

³³Ibid., p. 276.

³⁴Ibid., p. 347, n. 33.

³⁵Donnellan, K. "Speaking of Nothing" in Contemporary Research in Philosophical Logic and Linguistic Semantics, Hockney et. al. (eds.), 1973, pp. 93-118.

CHAPTER III

THE HISTORICAL EXPLANATION THEORY

INTRODUCTION: In the last chapter it was mentioned that Donnellan has defined a relationship that can hold between the user(s) of a name, the name and the thing named in virtue of which the thing is named and that does not involve the thing satisfying any description associated with the name. A name, according to this theory, will be a rigid designator since the meaning of the name will not be equivalent to the meaning of some particular description or cluster of descriptions. In this chapter we will examine the theory which Donnellan calls "the historical explanation theory". The historical explanation theory can be divided into two parts: the negative aspect and the positive aspect.

In the first section of this chapter we will argue that the negative aspect of the historical explanation theory is closely connected to the positive aspect of that theory. According to Donnellan, the negative aspect of the historical explanation theory, i.e., what the theory denies and its reasons for doing so have been better represented in the literature than the content of the positive theory. Donnellan mentions Kripke's article "Naming and Necessity" as one which gives arguments for the negative aspect of the theory. There are, however, important differences between the negative account of

the historical explanation theory that Kripke gives and Donnellan's own account of the negative aspect of that theory. The most important difference, as we will point out in the first section, is that Donnellan's counter-examples to the principle of identifying descriptions have an air of artificiality, whereas Kripke's do not. Thus, in the first section we will argue that Donnellan's counter-examples to the principle of identifying descriptions do not succeed in discrediting that principle unless they are modified considerably. The modified counter-examples will enable us to see more clearly why the negative aspect of the historical explanation theory has to be supplemented with the positive aspect. Finally, in the second section of this chapter we will examine the positive aspect of the historical explanation view.

I

THE NEGATIVE ASPECT OF THE HISTORICAL EXPLANATION THEORY

1. DONNELLAN'S ARGUMENTS AGAINST THE PRINCIPLE OF IDENTIFYING

DESCRIPTIONS: The negative aspect of the historical explanation theory consists in a criticism of the thesis which Donnellan calls "the principle of identifying descriptions". Donnellan formulates the principle as follows:

"It states, in the first place, that ... the user(s) of a proper name must be able to supply a set of ... 'non-question begging descriptions in answer to the question 'To whom (or what) does the name refer?' ... Secondly, the principle states that the referent of a proper name (as used by a speaker in some particular utterance), if there is one, is that object that uniquely fits a 'sufficient' number of the descriptions in the set of identifying descriptions. As a

corollary, when no entity (or more than one) satisfies this condition, the name has no referent, and a negative existential statement expressible by a sentence of the form 'N does not exist' (where 'N' is the name in question) will be true".¹

Donnellan's formulation of the principle is such as to accommodate, both, the Description theory and two versions of the Cluster theory of names, i.e., those of Searle and Strawson. Thus, the sufficient number of descriptions that an object must satisfy might be all the descriptions associated with the name as on the Frege-Russell view, or they might be an indeterminate number, as Searle and Strawson would claim. Donnellan proceeds to give counter-examples to the principle as it has been formulated by him. We have already dealt, in the last chapter, with the Description theory. So, before examining Donnellan's counter-examples to the principle of identifying descriptions, we will try and formulate the views of Searle and Strawson more precisely.

Since the principle of identifying descriptions, as it has been formulated by Donnellan, is a two-stage thesis, we will try and formulate the views of Searle and Strawson in two stages. The principle of identifying descriptions requires, first of all, that the speaker be able to supply a set of descriptions in answer to the question "To whom (or what) does the name refer?" Searle has the following to say about the first stage of the principle of identifying descriptions:

"..., anyone who uses a proper name must be prepared to substitute an identifying description (remembering that identifying descriptions include ostensive presentations) of the object referred to by a proper name. If he were unable to do this, we should say that he did not know whom or what he was talking about ..."²

Searle goes on to qualify this in order to accommodate certain counter-examples. For, suppose I learn the use of the name "Aristotle" by being told that it is the name of the Greek philosopher who taught Alexander. However, if I later learn from Greek scholars that Aristotle did not teach Alexander, I could not accuse them of self-contradiction which I would be forced to do if the description "the teacher of Alexander" was the only one which I associated with my use of the name "Aristotle". Thus, since a particular description may turn out to be false, Searle proposes to substitute a set of descriptions for just one description. Thus, he says:

"Suppose we ask the users of the name 'Aristotle' to state what they regard as certain essential and established facts about him. Their answers would constitute a set of identifying descriptions ..."³

The second part of the principle of identifying descriptions states that the referent of a proper name, if there is one, is that object that uniquely fits a "sufficient" number of the descriptions in the set of identifying descriptions. Once again Searle's Cluster theory of names satisfies this requirement.

"... suppose we have independent means of identifying an object, what then are the conditions under which I could say of the object, 'This is Aristotle?' I wish to claim that the condition, the descriptive power of the statement, is that a sufficient but so far unspecified number of these statements (or descriptions) are true of the object. In short, if none of the identifying descriptions believed to be true of some object by the users of the name of that object proved to be true of some independently located object, then that object could not be identical with the bearer of the name. It is a necessary condition for an object to be Aristotle that it satisfy at least some of these descriptions. This is another way of saying that the disjunction of these descriptions is analytically tied to the name 'Aristotle' ..."⁴

Thus, Searle's Cluster theory of names satisfies both the requirements which Donnellan has included in the principle of identifying descriptions.

Strawson's Cluster theory of names, too, satisfies both the requirements that have been stated by the principle of identifying descriptions. Thus, for example, Strawson writes:

"Consider the situation in which a reference is made, by name, to Socrates. ... both speaker and hearer, in this situation, satisfy the conditions for successful term-introduction if each knows some distinguishing fact or facts, not necessarily the same ones, about Socrates, facts which each is prepared to cite to indicate whom he now means, or understands, by 'Socrates'. But what is the relation between these facts and the name? Suppose we take a group of speakers who use, or think they use, the name, 'Socrates', with the same reference Now it would be too much to say that the success of term-introduction within the group by means of the name requires that there should exist just one person of whom all the propositions in the composite description are true. But it would not be too much to say that it requires that there should exist one and only one person of whom some reasonable proportion of these propositions is true."⁵

It is quite clear from the above passage that in order for a name to be introduced successfully the user(s) of the name must be able to supply a set of distinguishing facts about the referent of the name. And, secondly, the referent of the name is going to be that object of whom some reasonable number of descriptions in the set of identifying descriptions are true. Hence, both, the theories of Searle and Strawson satisfy the requirements that have been formulated by Donnellan in the principle of identifying descriptions. We will now examine Donnellan's objections to the principle of identifying descriptions.

As we have already seen, the first part of the principle of identifying descriptions states that the user(s) of a name must be able to supply a set of identifying descriptions. Donnellan constructs the following counter-example to show that the user(s) of a name need not supply a set of descriptions which identifies the referent of their use of the name.

"Suppose a child is gotten up from sleep at a party and introduced to someone as 'Tom', who then says a few words to the child. Later the child says to his parents 'Tom is a nice man'. The only thing he can say about 'Tom' is that Tom was at a party. Moreover, he is unable to recognize anyone as 'Tom' on subsequent occasions. His parents give lots of parties and they have numerous friends named 'Tom', ... Does this mean that there was no person to whom he was referring? It seems to me that his parents might perfectly well conjecture about the matter and come up with a reasonable argument showing that the child was talking about this person rather than that."⁶

Why would the parents of the child be inclined to say in this situation that their child has referred to someone? Probably because they think that their child is capable of meaning something by uttering the English sentence "Tom is a nice man". The parents thus take the trouble of trying to identify the referent of their son's use of the name "Tom" because they believe that their child intends to refer to someone, i.e., he has some particular person in mind. But then it is up to Donnellan to explain what it is to intend to refer to an object and succeed or fail in doing so. And Donnellan does, in fact, give an explanation of what it is to intend to refer to an object. We will deal with this when we examine his account of the distinction between the referential and attributive uses of definite descriptions. Yet, even without an account of what it is to intend to refer to an

object, we can see that Donnellan's counter-example shows that the requirement that every use of a name have a backing of descriptions, is highly dubious. However, of greater importance are the counter-examples to the second stage of the principle of identifying descriptions, and it is here that a certain amount of artificiality is introduced by Donnellan in connection with his counter-examples.

According to Donnellan, the artificiality of the counter-examples is forced on him by the Cluster theory of names as it has been formulated by Searle and Strawson, since this theory invokes the notion of an object satisfying a "sufficient" number of descriptions associated with a name in order to be regarded as the referent of the name. Now, the more artificial a counter-example, the more we have to wonder how much importance we are to attach to the counter-example. But this fact can be overlooked if, as Donnellan claims, the artificiality of the counter-examples is, indeed, forced on him. However, it is not clear that it is. Most of Kripke's examples against the Cluster theory of names can hardly be regarded as artificial and yet they serve to nail down the position against the Cluster theory of names beyond reasonable doubt. So, in the case of Donnellan's counter-examples the ingeniousness of the scenarios must be regarded as depriving them of their importance. But it is not very difficult to think of counter-examples to Searle's and Strawson's theories and we will construct two such counter-examples.

To construct counter-examples to the principle of identifying descriptions we have to show that there are possible situations in which the referent of a name does not fulfill the requirements

mentioned in the principle of identifying descriptions, or situations in which an object satisfies those requirements but is not the referent. We will consider, first of all, a situation in which an object satisfies the condition stated in the principle of identifying descriptions but is not the referent of a name. The counter-example is a modification of one of Donnellan's examples. Suppose, a teacher of philosophy who knows a great deal about Aristotle and his student who knows less than him about Aristotle, are conversing with each other about Aristotle. Suppose, furthermore, that the descriptions which the student associates with his use of the name "Aristotle" are derived from his teacher. Now, given the second part of the principle of identifying descriptions (henceforth ID), when the student and the teacher affirm the existence of Aristotle one could be right and the other wrong, since they both possess different criteria for identifying Aristotle. Furthermore, suppose that the description(s) which the student associated with his use of the name "Aristotle" are in fact true of Plato, whereas the descriptions which the teacher associated with his use of the name "Aristotle" are true of Aristotle. According to the second stage of ID we would have to say that the student used the name "Aristotle" to refer to Plato. But in the context of our scenario this is absurd. For the social context surrounding the utterance is such as to preclude the possibility of the referent of the name "Aristotle" turning out to be Plato. The context is one in which the student and the teacher are conversing together using the name "Aristotle". And in such a context it seems natural to ascribe to the student the intention to

refer to the same man to whom his teacher refers when he uses the name "Aristotle". If this is the student's intention then we can reject the conclusion that the student refers to Plato when he uses the name "Aristotle". Moreover, this rejection runs counter to ID. For, an advocate of ID cannot now say that the description that the student associated with his use of the name "Aristotle" is "the man referred to by my teacher", for that would be circular.

In constructing the second counter-example to ID we will try to show that there are possible situations in which the referent of a name does not fulfill the requirements mentioned in ID. Consider, for example, a name such as "Yevtushenko". Now, although some people may be able to give a description that picks out Yevtushenko uniquely, most of us when we use the name will say: well he's a poet or something. We may not think that this picks out anyone uniquely. But I still think that we use the name "Yevtushenko" as a name for Yevtushenko. Here, then, we have a case where we are not able to supply a backing of descriptions which will pick out Yevtushenko uniquely, but where we still use the name "Yevtushenko" as a name for Yevtushenko. In the next chapter we will mention other counter-examples to Searle's and Strawson's theory of names which are due to Kripke. But there is another way of seeing that the Cluster theory of names, as formulated by Searle and Strawson, must be false. To see why it must be false we have to examine Donnellan's distinction between the referential and attributive uses of definite descriptions. The reason for this is that descriptions are paradigms of referring expressions and a particular treatment of descriptions can lead one

to adopt a similar treatment of proper names. Thus, for example, Donnellan writes:

"The model referring expression has been for many philosophers of language, I believe, a definite description ... An object is referred to in virtue of possessing uniquely the properties mentioned in the definite description. It is not hard to see how this standard leads to adopting the principle of identifying descriptions for proper names. Proper names are referring expressions, yet on the surface fail to exhibit any descriptive content. Given definite descriptions as the paradigm, one is forced to look under the surface ... for 'the backing of descriptions' that must be there."⁷

2. DONNELLAN ON ATTRIBUTIVE AND REFERENTIAL DESCRIPTIONS: In his article "Reference and Definite Descriptions"⁸, Donnellan argues that the following kind of situation tells against Russell's theory of definite descriptions. Suppose someone at a party, looks in a certain direction and says to his companion

(1) "The man over there drinking champagne is happy tonight"

Suppose the man to whom both the speaker and the hearer refer is actually drinking water but he is, nevertheless, happy. Now, if there is no one drinking champagne over there, Russell would regard (1) as false. However, as Donnellan points out, with cases that are similar to (1) we think that the speaker said something true of the man to whom he referred in spite of his misimpression. Consider, once again, the following example. Someone sees a woman with a man. Taking the man to be her husband, and observing his attitude towards her, she says "Her husband is kind to her". Now, suppose that the man in question is not her husband. Suppose he is her lover, to whom she has been driven by her husband's cruelty. Once again Russell

would say that the statement is false, on the basis of the cruelty of a man no one was talking about.

To handle these cases Donnellan distinguishes two uses of definite descriptions. In the "attributive" use, a speaker "states something about whoever or whatever is the so-and-so".⁹ In the "referential" use, a speaker "uses the description to enable his audience to pick out whom or what he is talking about and states something about that person or thing. In the first (attributive) case, the definite description might be said to occur essentially, for the speaker wishes to assert something about whatever or whoever fits that description; but in the referential use the definite description is merely one tool for ... calling attention to a person or thing ... and ... any other device for doing the same job, another description or name, would do as well".¹⁰ We can illustrate the distinction by considering another example from Donnellan. Suppose I encounter the body of Smith who has been murdered. The condition of Smith's body leads me to say "Smith's murderer is insane". Here we have an attributive use: we are speaking of the murderer whoever he may be. On the other hand, suppose that Jones is on trial for Smith's murder. Observing the neurotic behaviour of the defendant, I may say "Smith's murderer is insane" (where I have forgotten the name of the defendant but am convinced of his guilt). Then my use is referential. One test for deciding whether a description is being used attributively is the appropriateness of the comment "whoever he is". Thus, in the first case we may say "Smith's murderer (whoever he is) is insane", but not in the second.

According to Donnellan, there are several things which may be regarded as being true of the referential use of a description but not true of the attributive use. Thus, if a speaker S uses a definite description "the ϕ " referentially then the speaker will at least intend that there should be an object about which the following may be regarded as true.¹¹ x is the intended object.

- (a) S will have referred to x whether or not x is the ϕ .
- (b) S will have said something true or false about x whether or not x is in fact the ϕ .
- (c) S, in using "the ϕ " to refer to x will have presupposed or implied that x is ϕ .
- (d) In reporting S's speech act we can use expressions other than "the ϕ " or expressions that are synonymous to "the ϕ ".

Of all these theses, (a) may be said to be the most controversial. It is, of course, true that a speaker may refer to an object x by means of the description "the ϕ " even though the object does not uniquely satisfy the property mentioned in the description. Thus, for example, when I use the description "the Great Emancipator", I could refer to Abraham Lincoln despite the fact that there were other great emancipators like Moses, Jesus, etc. The problem is, rather, can reference fail when a definite description is being used referentially. Let us examine Donnellan's view regarding this matter in "Reference and Definite Descriptions".

According to Donnellan if a speaker S intends to refer to an object then he does not fail to do so simply because his audience fails to identify the object which he has in mind. Moreover, the

speaker may refer even if nothing fits the description which is being used referentially. Donnellan goes even further and says that a speaker may succeed in referring to an object to which he intends to refer even if he does not believe that the description being used by him fits any object. This last requirement clearly conflicts with condition (c) which says that if a speaker uses the definite description "the ϕ " referentially to refer to an object then he presupposes or implies that the object is ϕ . However, if the speaker does not believe that the object he intends to refer to fits the description which the speaker uses with the intention to refer to the object, then how can the speaker be said to presuppose or imply that the object fits the description? Donnellan gets around this difficulty by admitting that those cases where a speaker does not believe that the description he uses fits any object, are not normal cases of the use of a description and are, in fact, "parasitic on a more normal use". Thus, he says:

"When the definite description is used referentially ... the presupposition or implication stems simply from the fact that normally a person tries to describe correctly what he wants to refer to because normally this is the best way to get his audience to recognize what he is referring to."¹²

For all practical purposes, then, we can ignore the case where a speaker does not believe that the description which he uses fits any object. This seems a sound policy given the fact that it is not even clear whether the speaker has referred in these cases. For, if a necessary condition for referring is that a speaker should intend to refer, then, clearly, that condition is not satisfied in the above

case. As a matter of fact, where the speaker does not believe that the description which he uses fits the object which he has in mind, he may be said to intentionally mis-refer to the object in question. This, in spite of the fact that he may succeed in communicating to his audience who or what he is talking about. To illustrate this, consider the following example from Donnellan. Suppose I wish to communicate to the followers of a man whom I believe to be a usurper and not a king. Now, the followers of this man may regard him as a king and not as a usurper. Furthermore, I may know that they thus regard him as a king. With this in mind I may utter the sentence "the king is holding court". It seems natural to say in this case that my primary intention is to communicate with the followers of the usurper whom I'm talking about. And in order to accomplish this I intentionally mis-refer to the usurper as the king. Anyway, not much hinges on it once we recognize that these cases are indeed parasitic on the normal uses of descriptions.

Let us return to our original question, which was whether one can fail to refer by means of a description if the description is being used referentially. We have seen that according to Donnellan it is not necessary that in order for us to succeed in the act of referring, the audience be able to identify the object to which we intend to refer. We can refer to an object even though the audience is unable to pick out the object to which we have referred. Moreover, the speaker may refer even if there is no unique object which fits the description being used by the speaker. Donnellan even allows for a referential use without a referent. This, presumably takes place

"when no entity can be correctly identified as 'what the speaker meant to be talking about', although the speaker intended that there should be. This 'failure of reference' will not come about simply because nothing is ϕ or because more than one thing is ϕ and, moreover, could happen when there is a unique ϕ ... exactly when it will happen could be specified only after an analysis of the criteria for deciding what entity it is to which the speaker is referring, an analysis I have not tried to give in the original paper".¹³ It is not too clear what Donnellan is saying here but it seems that he wants to distinguish between two senses of "reference". The sense of reference with which he is not concerned in the article from which the above passage has been quoted is one in which the speaker intends to refer to an object, in using a description referentially, and succeeds in doing so, even though there is no entity which can be correctly identified as the referent by the speakers audience. In this sense the speaker may succeed in referring to an object, but, because his audience is unable to identify the object to which he has referred, he fails to refer to it in another sense. Thus, the "failure of reference" does not consist in there not being an object but in the audiences' inability to identify the object of reference. The sense of reference which concerns Donnellan here is illustrated by him by means of the following example.

Suppose the speaker thinks that he sees a man at a distance. He may then ask the question "Is the man carrying a walking stick the professor of history?" Donnellan distinguishes three cases in which the speaker may be said to refer and one case in which it is

doubtful whether he could be said to refer. In the first case there is a man in the distance carrying a walking stick. Here the speaker may be said to have referred to the man and asked a question about him which can be answered given the appropriate information. In the second case the man is not carrying a walking stick but an umbrella. Here, too, the speaker may be said to refer to the man although someone may correct the speaker's misimpression that the man is carrying a walking stick. Once again the speaker's question may be answered by someone with the information at his disposal. In the third case there is no man in the distance but a rock which looks like a man. According to Donnellan, in this case, too, the speaker may be said to have referred to something although it is now not clear that the question he asked can be answered correctly. But the reason why the question cannot be answered is not because the speaker has not referred but because it would be misleading if someone who knew that the speaker had mistook the rock for a man, tried to answer the question. Finally, there is the case where there is nothing to which the speaker may be said to have referred. There is no rock or anything else which looks like a man carrying a walking stick. Thus, even though the speaker intended to refer to an object he fails to do so because there is nothing to which he may be said to have referred. In this case there is nothing of which the speaker can say "I was referring to that although I now see that I was mistaken in using the expression which I did use to refer to it". Thus, we finally have an account of what it is for a speaker to intend to refer to an object and fail in doing so. A speaker fails to refer to an object if there is nothing

to which he could be said to have referred even if mistakenly. In the next sub-section we will examine this view.

3. HUMPTY DUMPTY AND REFERRING: Mackay has criticized some of the things that Donnellan says about intentions by comparing his notion of intention to Humpty Dumpty's theory of meaning:

"There are three hundred and sixty-four days when you might get 'un-birthday presents.'
 'Certainly', said Alice.
 'And only one for birthday presents, you know. There's glory for you!'
 'I don't know what you mean by "glory",' Alice said. Humpty Dumpty smiled contemptuously. 'Of course you don't—till I tell you. I meant "there's a nice knockdown argument for you!"'
 'But "glory" doesn't mean "a nice knockdown argument",' Alice objected.
 'When I use a word', Humpty Dumpty said, in rather a scornful tone, 'it means just what I choose it to mean—neither more nor less'."¹⁴

According to Mackay, Donnellan is giving a Humpty Dumpty account of referring. The account he gives of referring is such that it leads to a collapse of referring into intending to refer. We have seen that this is not the case since a speaker can fail to refer when he is unable to identify the object to which he intended to refer. Let us examine Mackay's criticism more closely. In considering Donnellan's examples involving the referential use of descriptions, Mackay distinguishes four elements, namely: (a) the intentions of the speaker; (b) the referring expression; (c) the object to which the speaker intends to refer; and (d) the speaker's audience. Mackay understands Donnellan to be saying that if the audience is able to pick out the object to which the speaker intends to refer then the speaker may be said to refer and it doesn't really matter which referring expression

he has used. Thus, according to Mackay, the Humpty Dumpty view of referring that Donnellan is presenting is masked to a considerable extent by the fact that the examples which Donnellan gives of the referential use of descriptions involve "near misses". For example, he uses the description "the book on the table" when talking about the order "Bring me the book on the table". And he points out that it is possible for someone to fulfill this order even if there is no book on the table but one, say, besides it. Similarly, a speaker uses the description "the king" when talking about the usurper. Mackay thinks that if in these examples and others in which Donnellan illustrates the referential use of a description, Donnellan had used descriptions like "the rock" to refer to the book or "the Janitor" to refer to the usurper, then it would be much clearer to us why referring, on his account, collapses into intending to refer. Since the whole point of the referential use of a description is that a speaker may be said to have referred to an object even though the description does not fit the object, Mackay doesn't see any reason for not using outrageously inappropriate descriptions to talk about objects. But if one uses the expression "the rock" to refer to the book, then, according to Mackay, one could use any referring expression to refer to the book. Thus, referring expressions

drop out of the picture altogether and reference collapses into intending to refer.

Donnellan's reply to this criticism is quite straightforward. According to him, if it is merely a verbal dispute that what Mackay calls "what the speaker is talking about", Donnellan calls "what the speaker is referring to", then why not adopt Donnellan's way of talking? Mackay's reasons for not accepting Donnellan's way of talking have been mentioned above, namely, that they lead to a collapse of referring to intending to refer. Thus, Donnellan's reply to Mackay consists in refuting the charge that to refer is nothing more than to intend to refer. He does this by trying to show that it does not follow from the fact that in some cases a speaker can refer to a book by means of the description "the rock" that he could do this in any circumstance:

"The fact about intentions that I want to stress is that they are essentially connected with expectations. Ask someone to flap his arms with the intention of flying. In response he can certainly wave his arms up and down But this is not to do it with the intention of flying ... Perhaps one can, by a stretch of the imagination, conceive of someone (a child, say, who has seen birds flying) doing this. But such a person—the child, for example—would have expectations not shared by us ... To the next person who comes in the room I say, 'It's cold here'.

I have no expectations, any more than Humpty Dumpty did about Alice, that the person will construe my words in a novel way. Could I really intend that 'cold' should mean 'hot'? Or would my performance not be so much arm-flapping?"¹⁵

Thus, what the audience takes the speaker to be referring to depends on the expectations of the audience. We cannot say to the next man who comes into the room to bring us the rock on the table and succeed in referring to the book by means of "the rock" unless we believe that the person is going to take us to be referring to the book by the description "the rock". Most of the time even if our beliefs are false regarding what we take our description to be characterizing we succeed in referring to the object to which we intend to refer. The reason for this has been noted by Evans. The principle which our audience uses in these cases is one which "enjoins minimizing the attribution of inexplicable error."¹⁶

It is important to note that this is not the Principle of Charity which we will come across in the next chapter. It is with the intention to use an expression referentially to refer to an object in the presence of the audience with which we are here concerned. The principle of "minimization of inexplicable error" can be said to apply when the audience is capable of independently identifying the object. We can, perhaps, best explain this principle in the following way. When a person describes something, as when he describes what he is referring to, we are not limited to looking for something that fits his description (the Principle of Charity), or, rather, that fits his description better than anything else. We can also ask ourselves another question. We can ask "what thing would he

judge to fit these descriptions even if it does not really do so?"

In trying to answer this question the speaker's audience will utilize the descriptions which the speaker uses but the question will not be decided on the basis of what is denoted, if anything, uniquely by them. It is this question which we have to ask when dealing with the referential use of expressions. To sum up we can say that when a speaker intends to use an expression referentially he has certain expectations regarding his audience and regarding the object which he is attempting to characterize by means of the referring expression. These expectations consist in certain beliefs which the speaker holds about the audience and the object which the speaker attempts to describe. The speaker's beliefs may all be false or his expectations may not be fulfilled and the speaker may yet succeed in referring to the object to which he intends to refer. In these cases, the success of the act of reference would consist in the application of the principle of the "minimization of inexplicable error" by the speaker's audience.

We saw earlier on that Donnellan distinguishes between two senses of "reference" in characterizing the referential use of descriptions. In one of these senses the speaker intends to refer to an object and succeeds in doing so if the speaker's audience can correctly identify the object to which the speaker intended to refer by applying the principle of the "minimization of inexplicable error". This is the sense of reference we have been dealing with and since Mackay's criticism was directed against this notion of reference we can say that his criticism has been adequately dealt with. We will call this

notion of reference "social reference". But there is also the second notion of reference that Donnellan talks about in connection with the referential use of descriptions. Let us recall what this notion of reference was. According to Donnellan, a speaker may use a description or a proper name referentially intending to refer to an object and succeed in doing so even though there is no object which can be correctly identified as the referent by the speaker's audience. Let us call this notion "actual intended referent". Donnellan does not give a rationale for making a distinction between social reference and actual intended reference, but it is not too difficult to see why he would make this distinction. In "Reference and Definite Descriptions" and in "Putting Humpty Dumpty Together Again", the examples which Donnellan gives of the referential use of definite descriptions by a speaker are those where the speaker's audience is directly acquainted with the object of reference and is, therefore, in a position to correctly identify the object of reference.¹⁷ However, it is possible to use referring expressions referentially in those cases where the speaker's audience is not in a position to be directly acquainted with the object of reference, or, for that matter, neither is the speaker. Thus, for example, if I say to someone "the Prime Minister of Germany committed suicide in 1945", that person may reply, "But Germany did not have a Prime Minister in 1945". I could clarify the situation by saying something like "Oh, you know who I mean, I'm talking about Hitler". In this conversation I am willing to abandon the description when it becomes clear to me that the description cannot be used to achieve its purpose which was to use it to refer to

Hitler. Therefore, the description is being used referentially. But, at the same time, neither I nor the person I'm talking to are directly acquainted with Hitler or were directly acquainted with Hitler. Thus, the person I'm talking to cannot correctly identify the referent of my use of the description by applying the principle of "the minimization of inexplicable error". He has to do that by some other means. An even better example can be found in Kaplan:

"There was always something implausible about the idea that the referent of a proper name is determined - by the currently associated descriptions. For example, the entry under 'Ramses VIII' in the Concise Biographical Dictionary (concise Publications: Walla Walla, Washington) is 'One of a number of ancient pharaohs about whom nothing is known.'"¹⁸

It is possible to think of contexts in which someone uses the name "Ramses VIII" referentially without the audience being in a position to apply the principle of "the minimization of inexplicable error". Thus, for example, if I say to someone "it is possible that Ramses VIII was very rich" and he asks me who I'm talking about, I could reply by saying "One of a number of ancient pharaohs". Here the expression "One of a number of ancient pharaohs" does the job of calling attention to the person I'm talking about. However, my audience is hardly in a position to correctly identify who I am talking about, since the expression "One of a number of ancient pharaohs" does not pick out Ramses VIII uniquely and neither I nor my audience is directly acquainted with Ramses VIII. In this case, where the name is being used referentially, one could object that the speaker has not really referred to Ramses VIII although he has succeeded in communicating who it is that he is talking about. But it is difficult

to see how the speaker could have succeeded in communicating who it is that he is talking about since the audience cannot correctly identify who is being talked about, unless the audience takes the speaker to have referred to Ramses VIII. On the other hand, it would be wrong to argue that the speaker has not referred to anyone in using the name "Ramses VIII" because the name is being used attributively. For, as we have seen, an intuitive test for determining whether a referring expression is being used attributively is the legitimacy of the comment "whoever he is (was)", and this comment does not apply to the above scenario, although we can imagine a situation in which it does. For example, if I say to someone "One of a number of ancient pharaohs, it is possible, was very rich", and he asks me who I'm talking about, I might reply by saying "Oh, Ramses VIII, whoever he was". In this case it would be appropriate to say that the name "Ramses VIII" is being used attributively.

The difference, then, between the notion of social reference and that of actual intended reference is the following. In the case of social reference the speaker's audience is directly acquainted with the object to which the speaker intends to refer and, therefore, is in a position to independently identify the referent of the speaker's use of a referring expression by applying the principle of the "minimization of inexplicable error". Thus, if I order someone to bring me the book on the table then he can fulfill my order even if the book is not on the table but next to it, because he is directly acquainted with the object to which I intend to refer and can thus identify it by applying the principle which enjoins minimizing

inexplicable error. In the case of actual intended reference, on the other hand, no such means of identifying the object of reference are available, since neither the speaker nor the speaker's audience is directly acquainted with the object. Successful reference, in the case of actual intended reference will depend on other facts which we will examine in the next section.

To sum up. An examination of the referential use of definite descriptions has shown that we can use a definite description referentially to refer to an object even though the object does not uniquely possess the properties mentioned in the description. Since descriptions are paradigms of referring expressions we can see that treating proper names on the model of definite descriptions enables us to see more clearly why the Cluster theory of names, as formulated by Searle and Strawson, must be false. A proper name need not have a backing of descriptions, a sufficient number of which an object must satisfy, before being regarded as the referent of the name. For, just as the referential use of a definite description does not consist in the object referred to uniquely satisfying the property mentioned in the description, so the referential use of a proper name does not consist in the object referred to satisfying a sufficient number of the descriptions backing up the name.

II

THE POSITIVE ASPECT OF THE HISTORICAL EXPLANATION THEORY

1. ACTUAL INTENDED REFERENCE: In this section we will examine

Donnellan's account of what we are calling actual intended reference.

In presenting Donnellan's theory we will try and fill in the details since Donnellan only presents us with a sketch of his theory.

Donnellan formulates the positive aspect of the historical explanation theory in the following way.

"The main idea is that when a speaker uses a name intending to refer to an individual and predicate something of it, successful reference will occur when there is an individual that enters into the historically correct explanation of who it is that the speaker intended to predicate something of. That individual will then be the referent and the statement made will be true or false depending upon whether it has the property designated by the predicate."¹⁹

Donnellan gives an example to illustrate what is meant by "historical connection". If, for example, someone says "Socrates was snub-nosed" then the way to determine who he was referring to is by searching for an individual who is historically related to the speaker's use of the name "Socrates", and not by searching for an individual who best fits the descriptions associated by the speaker with his use of the name "Socrates". We search for an individual historically related to the speaker's use of the name. Donnellan introduces the notion of an omniscient observer of history to explain what is meant by an individual being historically related to a speaker's use of a name. Thus, he says;

"It might be that an omniscient observer of history would see an individual related to an author of dialogues, that one of the central characters of these dialogues was modeled upon that individual, that these dialogues have been handed down and that the speaker has read translations of them, that the speaker's now predicating snub-nosedness of something is explained by his having read those translations.

This is the sort of account that I have in mind by a 'historical explanation'."20

On this account, then, descriptions do play a role but not the one which was assigned to them by the principle of identifying descriptions. Consider, for example, the role of descriptions in connection with our use of proper names of historical figures. In most cases, according to the historical explanation theory, our use of proper names of persons with whom we are not directly acquainted is parasitic on uses of these names by other people. If we do associate a set of descriptions with names of historical figures, these descriptions would be derived from the things which other people have said about the presumed referent. For example, the descriptions which we associate with our use of the name "Socrates" could be traced back through many levels of parasitic derivation, to what was said using the name "Socrates", by Aristotle and Plato. But the descriptions which we might associate with our use of the name "Socrates", do not, on the historical explanation theory, play a role in determining the referent of the name. As Donnellan puts it

"... the historical explanation as seen by our omniscient observer may pick out an individual as the referent of the name 'Socrates' even though that individual is not correctly described by the speaker's attempt at identification. For example, the speaker may believe that Socrates, i.e., the person he refers to, was a philosopher who invented the Socratic method. But it is clearly imaginable that our omniscient observer sees that while the author of the dialogues did intend one of the characters to be taken as a portrayal of a real person, he modestly attributed to him a method that was his own brain child."21

Now, although the descriptions which we associate with a given name

do not determine the referent of the name, they can normally be regarded as reliable guides to the existence of a historical connection between our use of the name and the referent of the name. Most of the time, if the descriptions which a speaker associated with his use of the name "Socrates", for example, jibe more or less with the descriptions which we would associate with the name, then we simply assume that the speaker is referring to the same person to whom we refer when we use the name "Socrates". The ultimate test, however, for determining who the speaker is referring to is the existence of the appropriate sort of historical connection.

It is of some importance to note that the notion of an omniscient observer is a heuristic device which we can take quite seriously. Any theory of reference must allow for the possibility that its decipherment procedures are fallible. They could be wrong. The right decipherment is that of the omniscient observer. Thus, for example, there is a theory that Aristotle was not referring to the famous Socrates but rather to a snub-nosed individual of the same name who sat in the front row of his seminar.²² Hence, when a speaker asserts that Socrates was Snub-nosed, he may in fact be talking about Socrates the pupil of Aristotle whereas we all think that he is talking about Socrates-the-teacher of Plato. We may be wrong in thinking that the speaker is talking about the teacher of Plato. In this case the right answer, the right decipherment is that which an omniscient observer of history would provide. In other words, the omniscient observer plays the role which we cannot be expected to play. We cannot know in detail the history behind the uses of

names. For us the only reliable guides to the existence of a historical connection between a speaker's use of a name and the referent of a name are the descriptions which the speaker associates with his use of the name. We have to operate on the basis of the Principle of Charity which states that the referent of a speaker's use of a name is that object which best fits the descriptions which the speaker associates with his use of the name. We have nothing else to go on with in those cases where we are not directly acquainted with the referent of the speaker's use of the name.

The above account of the historical explanation theory, is, as Donnellan himself admits, only a sketch. But the theory is not so vague that we cannot see what Donnellan is getting at. Thus, we can offer the following as a first approximation of the type of conditions that an historical explanation theory of names might specify, and which have to be satisfied before we can say that a proper name is a name of such and such an object.

(H) If 'N' is a proper name uttered by a speaker S at a time t, then 'N' refers to x if and only if x is the one and only one object such that

(i) There is an appropriate historical connection between S's utterance of 'N' at t and the object x.

There are two points of clarification that ought to be mentioned concerning condition (H). The first is that clause (i) of condition (H) does not provide us with a formula for obtaining the referent of a name because it is not clear which historical connection is to be regarded as the "appropriate" one. The examples given by Donnellan

in this connection only show what is to count as an historical connection. But Donnellan has the following to say in support of his theory:

"In defence against this charge that the theory is excessively vague, it is helpful, I think, to compare it with another philosophical theory about a quite different problem. The causal theory of perception can be taken as holding that an observer, O, perceives an object, M, only if M causes O to have sense impressions. The theory seems to me to have content and to be important, whether or not it is correct. For one thing, if true it means that certain other theories are mistaken. But the theory as stated does not, obviously, allow us to say which among the various causal factors involved in an observer having sense impressions is the thing he perceives; nor does it tell us which ways of causing sense impressions are relevant. Possibly no philosophical analysis can determine this ... Analogously, the historical explanation theory lacks this sort of specificity. But for all that, if true, certain other theories, in particular the identifying descriptions theory, will be wrong and the theory does tell us something of importance."²³

The second point of clarification that ought to be mentioned is the following. Clause (i) of condition (H) provides us with a principle of decipherment for paradigmatic uses of proper names, i.e., uses of names like "Plato", "Kant", "Winston Churchill", ... etc. According to this view, a proper name such as "Plato" or "Kant", i.e., a name which is not referentially ambiguous, is deciphered on a particular occasion of its use in terms of the appropriate historical connection between the utterance of the name by the speaker and the object which bears the name. However, there are some proper names which are referentially ambiguous, i.e., names like "Zeno". When a speaker utters the name "Zeno", for example, we want to determine whether he is talking about Zeno the Stoic or Zeno the Eleatic. In other words,

we want to disambiguate his use of the name "Zeno", we want to determine which object the speaker has in mind in uttering the name "Zeno" at a given time. Thus, for referentially ambiguous names we have to add an extra clause to condition (H). We can re-write condition (H) to accommodate referentially ambiguous names in the following way:

(H) If 'N' is a proper name uttered by a speaker S at a time t, then 'N' refers to x if and only if x is the one and only one object such that

(i) There is an appropriate historical connection between S's utterance of 'N' at t and the object x.

(ii) S has x in mind in uttering 'N' at t.

Thus, a name which is ambiguous is to be deciphered on a particular occasion of its use by means of two facts: (a) which individual the speaker has in mind in uttering the name in question, and (b) whether or not there is an appropriate historical connection between the utterance of the name by the speaker on the occasion in question and the individual which the speaker has in mind.

To arrive at a final account of the historical explanation theory we have to examine Donnellan's account of the way in which the historical chain terminates and his account of the notion of a block. We will do this in the next sub-section.

2. TERMINATION AND BLOCKS: The notion of a block is introduced by Donnellan in connection with the problem of negative existence statements. So it will be worthwhile to briefly introduce the problem of negative existence statements.

The problem of negative existence statements can be briefly summarized in the following way. How is it possible to make a true statement about an object which does not exist? In order for a statement to be about something that thing must exist, otherwise how could the statement mention the thing or refer to it, for one cannot refer to nothing. It seems, therefore, that it is not possible to make true or false statements about an object which does not exist. We cannot even say that it does not exist. But, of course, we do make assertions in which we deny the existence of objects, such as "Santa Clause does not exist" or "Pegasus does not exist". And therein lies the puzzle. How is it possible for us to refer to objects only to say about them that they do not exist?

The question, then, which Donnellan attempts to answer is how the historical chain terminates when a speaker asserts a true negative existence statement. Consider, for example, the statement "Santa Claus does not exist". According to Donnellan, in order for this statement to be true a failure of reference would have to occur in other possible and actual predicative statements involving the name "Santa Claus". In other words, there will be no individual related historically to the use of the name "Santa Claus" in such statements as "Santa Claus comes at Christmas", etc. Thus, the question which has to be answered is what is involved in saying that a failure of reference occurs when proper names are used to make predicative statements. Donnellan introduces the notion of a block to explain failure of reference in predicative statements. Like the positive aspect of the historical explanation theory, the notion of a block

has not been too clearly formulated. But it is still possible to see what Donnellan is getting at. Basically, the idea is that if no individual can be identified as the referent of the use of the name, then we can say that a block has occurred in the history of the use of the name. It is not too clear what would count as identification in this case but we can at least point out what would not count as identification. Since the historical explanation theory denies that the referent of the speaker's use of the name is the object which satisfies a sufficient number of the descriptions which the speaker associated with his use of the name, we cannot say that the referent of a speaker's use of a name is identified by means of the descriptions which the speaker associates with his use of the name. Thus, the speaker must be capable of identifying the referent of the name independently of the descriptions which he associates with his use of the name. And when he is unable to do this we have an occurrence of what Donnellan calls a "block". For example, if a child who formerly believed in Snow White now finds out that Snow White does not exist, then we can say that what he has learned is that there is no individual historically related to his use of the name "Snow White" in predicative statements involving the name which he might have previously asserted. Rather, the historical claim connected with the child's use of the name in predicative statements ends in a block. The block in this case might be a story which has been told to the child as factual.

"... the block is the introduction of the name into the child's speech via a fiction told to him as reality by his parents. Blocks occur in other ways.

For example, children often invent imaginary companions whom they themselves come to speak of as actual. The block in such a case would occur at the point at which a name for the unreal companion gets introduced by the child himself via his mistaken belief that there is a companion to name."²⁴

Thus, one way in which the historical chain may terminate is via the notion of a block. But, there are two other ways in which a historical chain can terminate. We will consider these now.

Most of our beliefs about material objects will have their origin in a causal transaction which takes place between us and some object or the other. This causal transaction may take place or be derived through a long chain, from the transaction of others. Thus, we can form a belief about an object by means of perception (i.e., because we see it), or by reading about it, in which case our beliefs about the object can be traced back to the object via the author of the book or article. At this point we can ask how an object x , for example, can be the cause of the belief on the part of the speaker that $\phi(a)$? A rough answer would be the following. The object x is the cause of the belief on the part of the speaker, the belief which would be expressible in the form " $\phi(a)$ " if an event occurred in which the object x and the speaker were related causally in such a way that the belief that something ϕ 'd was produced. Now, in order for the name " a " to figure in the expression of the belief, one of the two possibilities have to be fulfilled. Either the name " a " must already have been connected by S with the object x , or in the situation described the speaker S bestows the name " a " on x . If the speaker has already baptized the object with the name then this only

takes us back one step. Thus, we are left with the situation where the object x causes the belief that something ϕ 's and S bestows the name on the object either by ostension or by means of a reference-fixing description. And if this happens we can say that a referential connection has been established between the object x and the name "a". So the two other ways in which a historical chain may terminate is if a referential connection has been established between a name and an object either by ostension or by means of a reference-fixing description.

There is one more issue which has to be resolved before we can write down the final version of the historical explanation theory. The issue revolves around the problem of deciphering a speaker's use of a name at a given time. For example, if a speaker denies the existence of Aristotle, we must have some means of determining whether he is referring to the philosopher or to the Greek shipping magnate. Domellan makes the following suggestion to deal with this problem:

"Certain uses of the name 'Aristotle' in predicative statements will have similar histories, histories that will distinguish them from other uses of the name. Each use of the name will, of course, have its own historical explanation, but these may, at a certain point, join up. So, in tracing back several uses of the name 'Aristotle' by me and several uses by you, we may find a common root in certain ancient writings and documents. While other uses of the name by me or by you may have nothing in common with the history of the first set of uses. It is possible that the histories may join at what I have called a block."²⁵

Hence, if a speaker uses the name "Aristotle" and we want to determine whether he is referring to the philosopher or the shipping magnate, we have to know something about the history of the use of the name, i.e.,

whether different uses of the name "Aristotle" have similar histories or not. If we incorporate Donnellan's suggestion we arrive at the following re-formulation of condition (H) which may stand as the final version of the historical explanation theory.

(H) If 'N' is a proper name uttered by a speaker S at a time t, then 'N' refers to x if and only if x is the one and only one object such that

(i) There is an appropriate historical connection between S's utterance of 'N' at t and x.

(ii) S has x in mind in uttering 'N' at t.

(iii) The historical connection between S's utterance of 'N' at t and x must be such that there is some characteristic which the historical chain possesses and by virtue of which it terminates at x

(iv) A historical chain may be said to terminate at x if and only if a referential connection has been established between a name and x either by ostension or by means of a reference-fixing description or via the notion of a block.

3. A PROBLEM: We have seen that the historical explanation of a speaker's use of a name must be taken into account in deciphering the speaker's use of the name. As we will see in the next chapter, the historical explanation theory does not differ from the causal theory of names in essential respects. Hence, the difficulties that arise in connection with the account which the historical explanation theory gives of the decipherment of names will be dealt with in the next chapter. We can however, briefly state the nature of the difficulties.

Consider, once again, the name "Santa Claus". Suppose a speaker utters the sentence "Santa Claus will come tonight" and we want to determine whether he is referring to the make-believe character or to one of his friends who is also called "Santa Claus". According to the historical explanation theory this will have to be done in the following way. We have to find a means of bringing together the uses of the name "Santa Claus" which refer to the make-believe character and distinguish them from other uses of the name which refer to the speaker's friend. The only way in which we can do this is by classifying those uses of the name "Santa Claus" which have similar histories from other uses of the name with different histories. But in order to do this we have to know some characteristic which the historical explanation of each use of name has and by virtue of which certain uses of the name "Santa Claus" can be distinguished from other uses of the name. Thus, we have to be able to trace back the various uses of the name until we find some feature which is common to certain uses of the name and which will distinguish them from other uses of the same name. But it is obvious that we do not and cannot be expected to do this, i.e., we do not and cannot be expected to trace back the historical chain connected with the speaker's use of the name "Santa Claus" before we can say one way or another who the speaker is referring to. We will consider this in greater detail in the next chapter when we deal with the causal theory of names.

SUMMARY: In this chapter we have dealt with the negative and the

positive aspects of the historical explanation theory. The important point which emerged from a discussion of the negative aspect of the historical explanation theory was that Searle's and Strawson's version of the Cluster theory of names was open to counter-examples. We gave two such counter-examples which discredit the principle of identifying descriptions. To further see that Searle's and Strawson's version of the Cluster theory of names is false, we were led to examine Donnellan's distinction between the referential and attributive uses of definite descriptions. In examining Donnellan's account of this distinction we saw that Donnellan distinguishes between two senses of reference, i.e., social reference and actual intended reference. We saw that in the case of social reference a speaker may be said to have failed to refer to the object to which he intended to refer just in case the speaker's audience is unable to identify the object of reference. To examine the notion of actual intended reference we had to turn to the positive aspect of the historical explanation theory which tries to preserve the rigidity of names by postulating a three-place relationship between the user of the name, the name and the thing named. It is a consequence of the historical explanation theory, as it has been formulated by us in condition (H), that in order for us to determine whom the speaker intends to refer to when he uses a name we have to know something about the historical explanation which lies behind the speaker's use of the name. But since we cannot be expected to know this it is clear that there is something wrong with the principle of decipherment provided by the historical explanation theory.

FOOTNOTES

¹Donnellan, K. "Proper Names and Identifying Descriptions" in Semantics of Natural Language, D. Davidson and G. Harman (eds.) (Reidel Publishing Company, Dordrecht, Holland: 1973), p. 360.

²Searle, J. "The problem of Proper names" in Semantics, Steinberg, D.D. & Jakobovits, L.A. (eds.) (Cambridge University Press; 1971), p. 137.

³Ibid., p. 138.

⁴Ibid., p. 138.

⁵Strawson, P.F. Individuals (Methuen & Co., London: 1964), p. 491.

⁶Donnellan, K. "Proper Names and Identifying Descriptions" in Semantics of Natural Language, op. cit., p. 364.

⁷Ibid., p. 358.

⁸Donnellan, K. "Reference and Definite Descriptions", The Philosophical Review 75 (1966), pp. 281-304.

⁹Ibid., p. 285.

¹⁰Ibid., p. 285.

¹¹See, for example, the following article by Donnellan: "Putting Humpty Dumpty Together Again", The Philosophical Review 77 (1968), pp. 203-215.

¹²Donnellan, K. "Reference and Definite Descriptions", op. cit., p. 292.

¹³Donnellan, K. "Putting Humpty Dumpty Together Again", op. cit., n. 6.

¹⁴Mackay, A. "Mr. Donnellan and Humpty Dumpty on Referring", The Philosophical Review, 77 (1968), p. 200.

¹⁵Donnellan, K. "Putting Humpty Dumpty Together Again", op. cit., p. 212.

¹⁶Evans, G. "The Causal Theory of Names", Proceedings of the Aristotelian Society, Suppl. vol. 47 (1973), p. 191.

¹⁷See, for example, "Reference and Definite Descriptions", op. cit., pp. 286-289.

¹⁸ Kaplan, D. "Bob and Carol and Ted and Alice" in Approaches to Natural Language, J. Hintikka and J. Moravcsik (eds.) (D. Reidel Publishing Company: 1973), p. 516, n. 8.

¹⁹ Donnellan, K. "Speaking of Nothing" in Contemporary Research in Philosophical Logic and Linguistic Semantics, Hockney, et. al. (eds.), pp. 93-118, cf. p. 104.

²⁰ Ibid., p. 104.

²¹ Ibid., pp. 104-105.

²² I owe this example to Dr. Wilson.

²³ Donnellan, K. "Speaking of Nothing" in Contemporary Research in Philosophical Logic and Linguistic Semantics, op. cit., pp. 106-107.

²⁴ Ibid., p. 110.

²⁵ Ibid., p. 112.

CHAPTER IV

CLUSTER THEORY OF NAMES

INTRODUCTION: In the last chapter we saw that the historical explanation theory tries to preserve the rigidity of a name by defining a relationship that can hold between the user(s) of a name, the name and the thing named in virtue of which the thing is named and that does not involve the thing satisfying any description associated with the name. But the historical explanation theory is not the only theory which defines a three-place relation between the user of a name, the name and the thing named. The causal theory of names also does that. In this chapter we will examine the question whether the relation which holds between the user of a name, the name and the thing named could be a causal relation. This chapter has two main objectives. The first task of this chapter is to expound the causal theory of names and to show that the theory is false. The second task of this chapter is to sketch a new theory of names which does justice to important intuitions which underlie the cluster theory of names and which, at the same time, preserves the rigidity of names. The causal theory of names has been put forward by Kripke as an alternative to the description and cluster theory of names. It was the outcome (i.e., the causal theory of names) of the difficulties, pointed out by Kripke, that beset the description and the cluster

theory of names, notably, the criticism that both the description and the cluster theories ignored the social character of naming. It is not our intention here to try and rehabilitate the description theory of names which we take to be false and against which we argued in some detail in the second chapter. However, since we will ultimately adopt a modified version of the cluster theory of names, we will start by giving an exposition of a particular version of the cluster theory of names (Section I); next, we will consider the difficulties raised by Kripke against the cluster theory of names and give an exposition of Kripke's causal theory of names (Section II); the difficulties involved in giving a causal account of names will be the subject matter of section III; and, finally, in the last section we will present the modified version of the cluster theory of names which we regard to be the correct one and which we take to preserve the rigidity of names.

I

THE CLUSTER THEORY OF NAMES

In the second and third chapters the cluster theories of names which were criticized were the ones defended by Wittgenstein, Searle and Strawson. It is not, however, with their version of the cluster theory of names with which we are concerned in this chapter. The cluster theory of names which we will present here is the one which has been proposed by N.L. Wilson.¹ Consider, for example, the name "Julius Caesar", as used by a person whom we will call "Charles". According to Wilson, the problem we are faced with is that of trying

to determine which individual (if any) is designated by the name "Julius Caesar" in Charles' language. Since in this context the words "designates" and "signifies" are used interchangeably by Wilson, we can say that the problem is one of trying to determine the significance of the name "Caesar" as it is used in Charles' language. According to Wilson, this question belongs to the field of theoretical descriptive semantics where we are concerned with giving truth-conditions for descriptive semantic statements. Descriptive semantics deals with factual questions such as "which language is the language used by Charles?" Thus, in trying to answer the question which individual (if any) is designated by the name "Caesar" in Charles' language, i.e., in trying to answer the question which belongs to the field of theoretical descriptive semantics we are laying down the procedure which the descriptive semanticist is to use in trying to determine the truth-value of factual statements. Wilson tries to show how descriptive semantics should be pursued when we are trying to determine which individual is designated by "Caesar" in Charles' language, by means of an example.

Suppose that Charles makes the following assertions which contain the name "Caesar".

- | | |
|--|------|
| (1) Caesar conquered Gaul. | (Gc) |
| (2) Caesar crossed the Rubicon. | (Rc) |
| (3) Caesar was murdered on the Ides of March. | (Mc) |
| (4) Caesar was addicted to the use of ablative absolute. | (Ac) |
| (5) Caesar was married to Boadicea | (Bc) |

According to Wilson the five statements taken together entail that

there is at least one individual who has the properties in question and at most one such individual. Thus, what Charles is really saying is:

(6) $E!(\forall x)(Gx \ \& \ Rx \ \& \ Mx \ \& \ Ax \ \& \ Bx)$

In (6) the name "Caesar" has been replaced by the variable "x" along with the E-Shriek-iota operator and (6) exhausts the cognitive content of the five statements made by Charles. Wilson concludes that proper names are really variables. However, in order to distinguish proper names from genuine variables, i.e., variables which occur in sentences like

(7) $(\exists x)(x \text{ disliked Caesar})$

Wilson proposes to call proper names quasi-variables.

"The difference between a genuine and a quasi-variable is this: when we use a genuine variable 'x' we close the scope of the quantifier quite promptly and it stays closed. When we use a quasi-variable or name, the scope of the initial E!-iota operator goes right down to the end and we add further asserted matrices as conjuncts to the scope of the operator as we come to have more knowledge. Thus, the totality of a person's beliefs is to be expressed in one vast multiply general sentence which changes and grows with the passage of time and the acquisition of more beliefs."²

Let us now try to determine the designation of the name "Caesar" in Charles' language. Given the five assertions that Charles has made about Caesar we conduct an empirical investigation and apply what Wilson has called the Principle of Charity. The designation of the name "Caesar" in Charles' language will be the individual which maximizes Charles' true beliefs, i.e., the individual which makes the largest possible number of Charles' statements true. Since the historical Julius Caesar satisfies more of the asserted matrices

which contain the word "Caesar" than any other individual, we can say that Caesar is the referent of the name "Caesar" in Charles' language.

Since the historical Julius Caesar is the designatum of the name "Caesar" in Charles' language, we can say that any asserted matrix such as "Caesar is F" in Charles' language is true if and only if Julius Caesar satisfies the matrix "x is F". Wilson call these primary truth-conditions. The sense of a proper name which is simply a function of the assertions which are made using the name corresponds to what Wilson calls secondary truth-conditions. The secondary truth-conditions for Charles' sentences about Caesar can be stated as follows:

"'Caesar is F' is true if there is one individual which satisfies more of the matrices 'Gx', 'Rx', 'Mx', 'Ax' and 'Bx' than any other individual does and that individual satisfies the matrix 'Fx'; and is false if there is one such individual and it does not satisfy the matrix 'Fx'; and is non-significant if there is no such individual."³

The primary truth-conditions are derived from the secondary truth-conditions and the secondary truth-conditions do not require that we conduct any empirical investigation before we can lay them down. A proper name can, thus, have a sense but no designatum. If, for example, we are willing to assert different things about Pegasus we endow the name "Pegasus" with a sense, but the name has no designatum or significance. Since sense is a function of the assertions which are made using the name and since different people may make different assertions using the same name, it follows that different people may use the same name with a different sense. Thus, if two people use the same name but attach different sense to the name, the success in communication would consist in their using the name with the same

significance:

"On this account it is significance rather than sense which offers itself as a replacement for meaning and 'significant' (rather than 'sensible') which offers itself as a replacement for 'meaningful' or 'having a truth-value'."⁴

To sum up we can say that an individual hooks on to a quasi-variable or an individual constant by virtue of the fact that it satisfies more asserted matrices in which the constant occurs than any other individual. In other words, the cluster of information that a speaker associates with his use of the name determines its reference by fit. If the speaker has no individuating information he will denote or refer to nothing.

II

TOWARD A CAUSAL THEORY OF NAMES

1. KRIPKE'S OBJECTIONS TO THE CLUSTER THEORY OF NAMES: In this subsection we will consider Kripke's objections to the cluster theory of names in some detail. Kripke states the cluster concept theory of names by means of a number of theses and a condition on the satisfaction of these theses.⁵ Accordingly, we will state the theses in the order in which he presents them and then consider the counter-examples he gives to these theses. The theses can be stated in the following way:

- (1) Associated with every designating expression 'X' there is a cluster of properties ϕ , such that S believes ' ϕx '.
- (2) One or some of the properties are believed by S to pick out some

individual uniquely.

(3) If most or a weighted most of the properties are satisfied by one unique object Y, then Y is the referent of 'X'.

(4) If no unique object is yielded by the vote then 'X' does not refer.

(5) It is known a priori by the speaker that if X exists, X has most of the ϕ 's.

(6) It is a necessary truth in the speaker's idiolect that if X exists, X has most of the ϕ 's.

(C) The account that a successful theory gives must not be circular. The properties which are used to pick out the individual must not themselves involve the notion of reference in a way that it is impossible to eliminate.

Some words of explanation are in order regarding Kripke's formulation of the above theses. Thesis (4) has to do with the voting procedure. As Kripke points out it is implausible that the most trivial properties are of equal weight with the most crucial. It is more plausible to suppose that some weighting takes place in these cases. The properties which are assigned the most weight are to be determined by some sort of vote. Theses (5) and (6) involve a distinction between a priori knowledge and necessary truth. According to Kripke, the notion of a priority belongs to epistemology and the notion of necessity belongs to metaphysics. His claim is that although it might turn out that everything a priori is necessary and vice versa, it will not turn out to be the case simply on the basis of definitional equivalence. It will require some philosophical argumentation. He

gives an example to show that there are necessary truths or falsehoods which are not known a priori:

"If the Goldbach conjecture is false, then there is an even number, n , greater than 2, and for no primes P_1 and P_2 , both $< n$, does $n = P_1 + P_2$. This fact about n , if true, is verifiable by direct computation, and thus is necessary if the results of arithmetical computations are necessary. On the other hand, if the conjecture is true, then every even number exceeding 2 is the sum of two primes. Could it then be the case that, although in fact every even number is the sum of two primes, there might have been an even number which was not the sum of two primes? What would that mean? Such a number would have to be one of 4, 6, 8, 10, ...; and, by hypothesis, since we are assuming Goldbach's conjecture to be true, each of these can be shown, by direct computation, to be the sum of two primes. Goldbach's conjecture, then, cannot be contingently true or false; whatever truth-value it has belongs to it by necessity ... But ... we don't know whether Goldbach's conjecture is true or false. So right now we certainly don't know anything a priori about it."⁶

The point of this example is that there are necessary statements which are not known a priori, and, thus, that the terms "necessary" and "a priori" are not synonymous, at least, on the face of it. A philosophical argument is required to show that they are synonymous.

Similarly, although it may be known a priori that if X exists, X has most of the \emptyset 's, it does not follow that it is a necessary truth that if X exists, X has most of the \emptyset 's. If, for example, a man has fixed the reference of the name "Aristotle" by means of the description "the teacher of Alexander", then in some sense he knows a priori that Aristotle is the teacher of Alexander. However, it will not be a necessary truth for the speaker that Aristotle is the teacher of Alexander. Aristotle might never have gone into pedagogy. So thesis (5) and (6) do not

mutually entail one another, and a counter-example to thesis (6) will not necessarily be a counter-example to thesis (5). Theses (5) and (6) also have converses which are derivable from theses (1) - (4). Thus, the converse of thesis (5) is that the statement "if there is some unique object which has most of the \emptyset 's in the properly weighted sense, then it is X" is known a priori by the speaker. And the converse of thesis (6) is that the statement "if there is some unique object which has most of the \emptyset 's in the properly weighted sense, then it is X" is a necessary truth in the idiolect of the speaker. Taking theses (5) and (6) together with their converses we can say that it is both a priori and necessary that something is X if and only if it has uniquely most of the \emptyset 's in the properly weighted sense. Let us now consider the counter-examples which Kripke gives to these (1) - (6).

Thesis (1), as Kripke points out, is a definition. So it can be regarded as true. Theses (2) and (3), according to Kripke, are both false and Kripke gives counter-examples to both these theses. Let us consider theses (2) first. The counter-examples that Kripke gives to thesis (2) are meant to show that either the uniqueness condition is not satisfied and yet the speaker may be said to refer, or the circularity condition is violated without preventing the speaker from picking out the referent of the name. Consider an example where the uniqueness condition is not satisfied. An ordinary man in the street may use the name "Feynman". When asked who he is talking about he may say "a famous physicist". Clearly, he does not pick out anyone uniquely and yet Kripke maintains that he still uses "Feynman" as

a name for Feynman. Next, consider a case where the circularity condition is violated. Suppose (once again) that the man on the street wants to refer to Einstein and he does this by using the description "the man who discovered the theory of relativity". Here he has given a unique description which is supposed to pick out Einstein. But if we ask him what the theory of relativity was he might reply "it was the theory that Einstein discovered". Here he has clearly violated the non-circularity condition. Now, Kripke wants to maintain that he has even in this case used the name "Einstein" as a name of the man Einstein. In other words, he has referred successfully to Einstein.

According to Kripke, similar counter-examples can be constructed to show the falsity of thesis (3) since it, too, depends on our being able to state unique conditions which are satisfied by an object. Thesis (3) states that if most or a weighted most of the properties are satisfied by a unique object Y, then Y is the referent of 'X'. Kripke gives an example to show that although most of the ϕ 's are in fact satisfied by a unique object, that object is not necessarily the referent of 'X' for the speaker. A speaker who associated with the name "Gödel" merely the description "the prover of the incompleteness of arithmetic" would nonetheless be referring to 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 stolen. If we agree that the speaker does not refer to Schmidt, then thesis (3) does not provide us with sufficient conditions for successful

reference to take place. If it is agreed that the speaker refers to Gödel then the conditions stated in thesis (3) are not necessary. Another counter-example to thesis (3) which Kripke gives is the following:

"What many people ... may 'know' about Peano is that he was the discoverer of certain axioms which characterize the sequence of natural numbers, the so-called 'Peano axioms'. ... these axioms are not actually due to Peano but to Dedekind ... So on the theory in question the term 'Peano', as we use it, really refers to ... Dedekind."⁷

Kripke's point is that if it is true that if most of the \emptyset 's are satisfied by a unique object Y then Y is the referent of the name 'X' for the speaker, then a speaker who only knew that Peano was the discoverer of certain axioms would be referring to Dedekind when he used the name "Peano". But since he doesn't it follows that thesis (3) is false. In a footnote to the article "Naming and Necessity", Kripke concedes that there are other uses of proper names like "Gödel" and "Peano", which satisfy the cluster theory of names.⁸

When someone says that Gödel proved the incompleteness of arithmetic then he may be taken to be referring to Gödel and not to Schmidt. But when someone says "Gödel relied on a diagonal argument in this step of the proof", he may be said to be referring to whoever proved the theorem. Following Donnellan this can be called the "attributive" use of proper names. On this account the sentence "Gödel proved the incompleteness of arithmetic" would be false given that it was Schmidt who actually proved the incompleteness of arithmetic. But the sentence "Gödel used the diagonal argument" would be true for some uses of the name "Gödel". The cluster theory of names, according to Kripke,

would still, in general, be false. But, as he admits, it would be applicable in a wider number of cases than he previously thought.

What about thesis (4) which asserts that if no unique object is yielded by the vote then "X" does not refer. This thesis has been shown to be false by one of the previous counter-examples, namely, the Feynman example. In the case of that example the vote did not yield a unique object but the speaker still used the name "Feynman" as a name of the man Feynman. But even if the vote yields no unique object or no object at all which satisfies a weighted number of the \emptyset 's, the speaker may still use a name to refer to an object. Thus, for example, though most people believe (falsely) that Einstein invented the atom bomb, no one person could really be said to have invented the atom bomb. And yet a speaker still uses the name "Einstein" as a name of the man Einstein when he says "Einstein invented the atom bomb".

Thesis (6) states "the statement 'if x exists, then x has most of the \emptyset 's' expresses a necessary truth (in the idiolect of the speaker)."⁹ We have already argued against this thesis in the second chapter, but let us recollect. Thesis (6) states that Aristotle, for example, has the logical sum of properties which are commonly attributed to him and that it is a necessary fact that he has these properties. As against this Kripke argues that most of the properties which are commonly attributed to Aristotle are properties which Aristotle might not have possessed. In a situation in which Aristotle did not do any of the things which he is commonly said to have done, we would describe that as a situation in which Aristotle did not do

any of those things. Similarly, in the case of a name like "Hitler":

"When I hear the name 'Hitler', I do feel its sort of analytic that that man was evil. But really, probably not. Hitler might have spent all his days in quiet in Linz. In that case we would not say that then this man would not have been Hitler, for we use the name 'Hitler' just as the name of that man, even describing possible worlds. (This is the notion which I called a rigid designator in the previous talks)!"¹⁰

Given that this is the case we must cross off thesis (6) as false.

But it is important to note that thesis (6) is not a part of Wilson's theory. It is Searle's cluster theory of names which is committed to thesis (6), whereas Wilson's theory is not. As we will point out when we are dealing with the amended version of the cluster theory of names, one can maintain that Aristotle is the person who in fact has a plurality of the properties which are commonly attributed to him, but he might have had almost none of them.

Thesis (5) which states that it is known a priori by the speaker that if x exists, x has most of the ϕ 's, has nothing to do with necessity and it can survive. If, for example, one fixes the reference of the name "Aristotle" as the man who did most of the things which are commonly attributed to him, then it will not be necessary that Aristotle did any of the things which he is claimed to have done. But it may be a priori in that this is how the reference of the name "Aristotle" is determined. So thesis (5) survives. The remaining theses, according to Kripke are all false. Theses (2) - (4) have counter instances. Thesis (6) has the consequence that names are not rigid designators and is to be rejected on those grounds alone. For, if it is a necessary truth that Aristotle did most of the things which he is commonly

claimed to have done, this can only be so if the meaning of the name "Aristotle" is equivalent to the meaning of a cluster of descriptions. But, in that case, as we have seen, the name "Aristotle" will not necessarily be a rigid designator.

According to Kripke, the picture which gives rise to the cluster theory of names is one which ignores the social character of naming. One thinks of the semantical subject as isolated from the rest of the community when determining the reference of a name such as "Gödel". Thus, a speaker might simply say that by "Gödel", I shall mean the man who proved the incompleteness of arithmetic. If a speaker does this and if it is actually Schmidt who proved the incompleteness of arithmetic, then when the speaker uses the name "Gödel" he is actually referring to Schmidt. But more important such an account of reference would run into a circle. If everyone determines the reference of the name "Gödel" by saying to himself that Gödel is the man who proved the incompleteness of arithmetic, then "none of us will get started with any attribution unless there is some independent criterion for the reference of the name other than 'the man to whom the incompleteness of arithmetic is commonly attributed'. Otherwise all we will be saying is, 'we attribute this achievement to the man to whom we attribute it', without saying who that man is, without giving any independent criterion of the reference ..."¹¹ Such an account would violate what Kripke has called the circularity condition (C). And no theory of reference should be circular.

In the last section we will examine Kripke's counterinstances to theses (1) - (4), but now let us ask ourselves what theory of names

Kripke offers as an alternative to the cluster theory of names. Kripke only presents a sketch of his theory and we will have to fill in the details. The theory that Kripke presents might best be characterized as the causal theory of names.

2. THE CAUSAL THEORY OF NAMES: A causal theory of reference like any other theory of reference, must provide a set of conditions which have to be satisfied before we can say that a given proper name N is a name of such and such an object X. Kripke does not claim that he is giving a set of necessary and sufficient conditions which have to be satisfied in order for a name to be a name of an object. But he does claim that he is giving a better and quite different picture from the one given by the cluster theory of names. What is the picture that Kripke gives? According to Kripke, first of all an initial baptism takes place, i.e., the name is introduced by ostension, or by means of a definite description that fixes its reference. Subsequently, the name is passed on to others:

"When the name is 'passed from link to link', the receiver of the name must, I think, intend when he learns it to use it with the same reference as the man from whom he heard it. If I hear the name 'Napoleon' and decide it would be a nice name for my pet aardvark, I do not satisfy this condition. (Perhaps it is some such failure to keep the reference fixed which accounts for the divergence of present uses of 'Santa Claus' from the alleged original use.)"¹²

According to Kripke, then, once an object has been endowed with a given name, the name is passed on to others. This process continues as the name is passed from link to link of a chain of communication. What joins each link to the next is its causal connection with it,

together with the intention to use the name with the same reference as the previous speaker. One can think of the intermediate links in such causal chains of communication as points at which one speaker acquires a way of using the name from another person, by witnessing the others use and by forming an intention to use the name to refer to the same thing as does the other person. Consider, for example, the following situation. A kitten is acquired by a man (let us call him "Kuba") and his wife. There is an initial baptismal ceremony in which the kitten is given the name "Ginger", by Kuba and his wife. The two of them gain their ability to refer to Ginger at the naming ceremony. Everyone else will borrow their reference either directly or indirectly from these two. Thus, someone else might be introduced to the kitten and told "This is Ginger". And this person in turn might tell someone else that Kuba's cat is called "Ginger". So the name will be passed on from one person to another. The acquisition of the name by one person from another would constitute part of the social practice involving the name. And underlying the use of the name "Ginger" will be causal chains grounded in Ginger. Thus, the fact that a speaker in a community is able to use a name to refer to the same object as the other speaker's in the community, would be accounted for in terms of the causal chain extending all the way back to the initial baptismal ceremony.

For the time being let us ignore the bit about intentions by offering the following as a first approximation of the type of condition that a causal theory of names might specify as constituting the naming relation. Our reasons for not introducing the notion of an

intention at this stage will become clear as we proceed.

- (I¹) If "N" is a proper name uttered by a speaker S at a time t, then "N" refers to x if and only if x is the one and only one object such that:
- (i) S's utterance of "N" at t is causally connected with the use of the name by a group of speakers who believe that they are using the name with the same reference.
 - (ii) There is a causal connection between S's utterance of the name "N" at t and the object x.

The causal theory of names, as it has been formulated by us in condition (I¹), is meant to deal with the paradigmatic uses of proper names, i.e., uses of names like "Plato", "Gödel", "Winston Churchill", etc. These names do not require the notion of an intention and should be dealt with separately. Kripke, on the other hand, wants to introduce the notion of an intention in connection with the use of all proper names. We will point out the difficulty with this view below. But, first of all, let us try and determine Kripke's reason's for introducing the notion of an intention in connection with the use of all names.

There seem to be two reasons why Kripke introduces the notion of intending to refer to the same object as the person from whom one hears it (i.e., the name). First of all, he wants to eliminate those cases where someone uses the name "Napoleon", for example, which he has heard from another person, as a name for his pet aardvark. Secondly, the notion of an intention is introduced to account for reference shifts. Gareth Evans has given one such example involving a reference

shift:

"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."¹³

In this case, as Kripke points out, the present intention to refer to the African island overrides the previous intention to "preserve reference in the historical chain of transmission".

Now, there are cases where it makes sense to say that the speaker intends to refer to an object. Thus, for example, if I use the name "Pierre Trudeau" and intend by it to refer to my next door neighbour who is called "Pierre Trudeau" then I could be taken to have done so. But this is only possible because I can describe what it is to intend to refer to Pierre Trudeau, i.e., my next door neighbour rather than the liberal politician by saying something like "Pierre Trudeau—and I don't mean the liberal politician—did so-and-so". That is, it only makes sense to say that I intend to refer to Pierre Trudeau, my next door neighbour because I can describe what it is to refer to my neighbour independently of the existence of the intention to do so. It is because we know what it is to refer to an object independently of the presence of any intention, that it makes sense to say that in the case of ambiguity, as in the case of a name like "Madagascar" or "Zeno", it is the intention of the speaker which determines which object the speaker is referring to. Similar considerations apply in the case of Kripke's Napoleon example. Ordinarily we do not form an intention to imitate our informants. We just do it automatically. We can say that we refrain from using the

name in some new way, and not that we form an intention to use it in the same way as the person from whom we hear the name. To sum up we can say that condition (I¹) is to be regarded as constituting the naming relation, on the causal theory of names, for the paradigmatic uses of proper names. We will now try to emend condition (I¹) to accomodate those names which are referentially ambiguous. However, it is of some importance to separate condition (I¹) from its subsequent modifications which are meant to accomodate the use of names which are referentially ambiguous. For, one can argue that the causal theory of names is false because of its treatment of names which are ambiguous. But even if this is the case there will still be a motivation for holding the causal theory of names, since it may still be necessary to give a causal account of the use of a name on a particular occasion by means of the chains of communication.

3. REFERENTIAL AMBIGUITY: To deal with cases of referential ambiguity we have to introduce the notion of intention. There seem to be two intentions associated with the use of a referentially ambiguous proper name, on the causal theory of names. There is the intention to use a proper name to refer to the same individual as the person from whom one heard the name. And there is ~~also~~ the intention to refer to a pre-conceived object, i.e., the individual who was baptized with the name. This is borne out by the account which Kaplan gives of the causal theory of names.¹⁴

As Kaplan sees it, on the causal theory of names there are two critical intentions associated with the use of a proper name, namely,

the two intentions mentioned above. According to Kaplan, most uses of proper names involve a mixture of the intention to conform to a prior usage of the name, together with the intention to refer to a pre-conceived object. Thus, for example, if I introduce a name into your idiolect by means of a false introduction ("This is Jim O'Connor", but actually it isn't), you will be left with a mixture of the intention to refer to Jim O'Connor and the intention to refer to the man to whom you were introduced. We have already seen that it only makes sense to invoke the notion of intention in the case of referentially ambiguous names: Kaplan's example illustrates this point well because cases of misidentification are rare. However, there seems to be little doubt that for the causal theorist there will be two intentions associated with the use of referentially ambiguous names. For, the intention to use a name to conform to a prior usage does not eliminate the notion of reference. If, for example, I use the name "N" to refer to what X referred to by "N" who in turn referred to what Y referred to by "N" who ... etc. then none of us will get started unless there is some independent means of determining which object is the referent of the name "N": Thus, I must also intend to use the name "N" to refer to the object which was baptized with the name or the object with which I am acquainted, i.e., must intend to refer to a pre-conceived individual.

One way in which we might account for these mixed intentions is by distinguishing between speaker's reference and semantic reference. Following Kripke, we may tentatively define the speaker's referent of a designator to be that object which the speaker intends

to refer to on a given occasion and which he believes fulfills the conditions for being the semantic referent of a designator. The speaker uses the designator to make an assertion about the object in question which may not really be the semantic referent, if the speaker's beliefs that it fulfills the appropriate semantic conditions are in error. The speaker's referent is the thing or object the speaker referred to by the designator, though it may not be the semantic referent of the designator. Consider, for example, the following situation:

"Two men glimpse someone at a distance and think they recognize him as Jones. 'What is Jones doing?' 'Raking the leaves'. If the distant leaf-raker is actually Smith, then in some sense they are referring to Smith, even though they both use 'Jones' as a name of Jones."¹⁵

We don't have to agree with Kripke that the speaker's are referring to Smith when they use the name "Jones". In order to determine who it is that the speakers are referring to we have to determine which of the two of the speaker's intentions is the dominant one. We will deal with this below. For present purposes, we can say that if the speaker's have referred to Smith with the name "Jones", then the speaker's referent in this case is Smith, but Smith is not the referent of the designator "Jones" which the speaker uses to refer to Smith.

Thus, we might try to incorporate the notion of speaker's reference and that of semantic reference by offering the following as a first approximation of the type of conditions that a causal theory of names might specify as constituting the naming relation for referentially ambiguous names.

(I'') If "N" is a proper name uttered by a speaker S at a time t, then "N" refers to x if and only if x is the one and only one object such that:

(i) S's utterance of "N" at t is causally connected with the use of the name by a group of speakers who believe and intend that they are using the name with the same reference.

(ii) There is a causal connection between S's utterance of "N" at t and the object x.

(iii) S has x in mind in uttering "N" at t.

The semantic reference of a name "N" as uttered by a speaker S is determined by the causal connection between S's utterance of the name and the use of the name by a group of speakers who believe and intend that they are using the name with the same reference. The speaker's reference of the name "N" is determined in terms of the object which the speaker has in mind in uttering "N". For the causal theorist something like clause (iii) will be invoked when we try to decipher a name on a particular occasion of its use. Which individual is to be regarded as the referent of a referentially ambiguous name on a particular occasion of its use by a speaker is going to be determined by determining which object the speaker has in mind on the occasion in question. However, the satisfaction of clause (iii) cannot be sufficient for an object to be the referent of the use of a name.

According to the causal theory of names, there must also be a causal connection between the speaker's use of a name and the referent of the name. Therefore, we have to add the further requirement that the use of the name by a speaker on a particular occasion must be causally

connected with the use of the name by a group of speakers who believe and intend that they are using the name with the same reference, that is, clause (i) of condition (I"). But, since clause (i) does not eliminate the notion of reference, we have to add the further requirement that the reference of a name being used is determined by a causal chain of communication reaching back from the use of the name by a speaker to the referent of the name, i.e., clause (ii) of condition (I").

III

THE CAUSAL THEORY OF NAMES

1. OBJECTION TO THE ALLEGED SUFFICIENCY OF CONDITION (I"): There are several objections to the causal theory of names as it has been formulated by us in condition (I"). The first difficulty which confronts the causal theory of names has to do with clause (iii) of condition (I") which is meant to provide a principle of decipherment for the use of referentially ambiguous names by speakers. To see why it does not provide a correct principle of decipherment for names we only have to consider the following situation.

Two speaker's use the name "Jones" to refer to the man in the distance who is raking the leaves. Now, the man in the distance raking the leaves is actually Smith. Clearly, the two speakers intend to refer to Smith with the name "Jones", since they intend to refer to the man in the distance raking the leaves and that man is Smith. They also, of course, intend to refer to Jones, i.e., they intend to

use the name "Jones" as a name of the man Jones. Thus, the speakers have both Smith and Jones "in mind" when they use the name "Jones" to refer to the man in the distance raking the leaves. Now, on the account we have given of what it is for a speaker to have an object "in mind" we have to say that the two objects which the speakers have "in mind" are linked by means of two causal chains with two persons. But the speaker's use of the name "Jones" cannot refer to two persons. The problem remains as to who the speaker's are referring to when they use the name "Jones".

However, an extra clause can be added to condition (I'') to handle the above difficulty. We can say that the question whether the speakers are referring to Jones or Smith when they utter the name "Jones" can be determined by adding the further specification that one of the intentions, (i.e., either the intention to use a name in conformity with the use of the name by the other speakers in the linguistic community or the intention to refer to a pre-conceived object) is the dominant one. Kaplan, for example, makes this suggestion.¹⁶ In most cases the context will determine which of the two intentions is the dominant one. In those few cases where the contextual features are insufficient to determine which intention dominates, we may simply ask the speaker. So condition (I'') is now to be re-written as follows:

(I''') If "N" is a proper name uttered by a speaker S at a time t, then "N" refers to x if and only if x is the one and only one object such that

(1) S's utterance of "N" at t is causally connected with the use

of the name by a group of speakers who believe and intend that they are using the name with the same reference.

- (ii) There is a causal connection between S's utterance of "N" at t and the object x.
- (iii) S has x in mind in uttering "N" at t.
- (iv) S has x in mind because either the dominant intention is the intention to use the name "N" to refer to the object x which is the object referred to by the group of speakers with whose use of the name "N" S's utterance of "N" at t is causally connected; or because the dominant intention is the intention to refer to a pre-conceived object x.

⁶
 2. INTENTIONAL REFERENCE: We have seen that on the causal theory of names, the speaker's intention to refer to an object x will be determined in terms of the causal connection between the intention of the speaker to refer to x and the object x. So causality carries the chief burden in this attempted reduction of intentionality. There is much to get clear about the nature of this causal connection, but one thing is clear. The causal connection which holds between the speaker's state of mind and an object must be a reference-preserving connection. However, in order for us to decipher a speaker's use of a referentially ambiguous name it is not enough that there is an appropriate causal connection between the speaker's intention to refer to an object by means of the name and the object which is the referent of the name. We also need to know something about the causal chain by virtue of which the intention of a speaker is the intention to refer

to a given object which is the referent of the name in question.

The reason for this is the following.

Suppose a speaker intends to refer to Zeno the Stoic. We can characterize the speaker's intention to refer to Zeno the Stoic by means of the following sentence:

- (1) S intends that he refer with "Zeno" to Zeno.

The question which we have to answer is how we would distinguish the speaker's intention to refer to Zeno the Stoic from the speaker's intention to refer to Zeno the Eleatic which can also be characterized by means of (1). We cannot characterize the speaker's intention to refer to Zeno the Stoic as opposed to Zeno the Eleatic by invoking the notion of identification by description, i.e., we cannot characterize the speaker's intention to refer to Zeno the Stoic by means of the following sentence:

- (2) S intends that he refer with "Zeno" to Zeno the Stoic.

For, on the causal theory of names when the use of a name refers to an object, it would have referred to the object even if that object had failed to uniquely possess all or most of the properties which the speaker believed are uniquely possessed by the referent of his use of the name. But this means that a causal theorist cannot appeal to a speaker's intentions to refer to an object as the means whereby the speaker's use of the name is deciphered unless the intentions in question are other than those whose content can be given by means of an identifying description.

It seems, then, that the only way in which the causal theorist can distinguish between the speaker's intention to use the name "Zeno"

to refer to the Stoic, from his intention to use the name "Zeno" to refer to the Eleatic, is by distinguishing between the causal chains which determine the referents for the speaker's use of "Zeno" in the two cases. And in order to distinguish between the two chains he would need to know some characteristic which one chain possesses and the other lacks. The only possibility, then, seems to be that the characteristic possessed by the causal chain which determines the referent of the name "Zeno", where the speaker intends to refer to the Eleatic, is different from the characteristic possessed by the chain which determines the referent of the name "Zeno" in the case where the speaker intends to refer to the Stoic. Let us call the two characteristics " C_1 " and " C_2 ", respectively. We can now say that in order for the speaker to refer to Zeno the Eleatic there must be an appropriate causal chain possessing the characteristic C_1 between the speaker's state of mind and Zeno the Eleatic. And in order for the speaker to refer to Zeno the Stoic there must be a causal chain possessing the characteristic C_2 between the speaker's state of mind and Zeno the Stoic. It is important to note that the characteristics C_1 and C_2 do not function like Frege's notion of sense as Kripke thought.¹⁷ The speaker does not have to know what characteristic is possessed by the causal chain in order to know that he intends to refer to Zeno the Stoic rather than Zeno the Eleatic. It is we who have to know something about the chain in order to decipher the speaker's use of the name. But, of course, if these characteristics functioned like Frege's notion of sense then the speaker would have to grasp or understand that the chain in question possesses one of

these characteristics before he can say that he uses the name "Zeno", for example, to refer to Zeno the Eleatic. At this point we can re-write condition (I^{'''}) as follows:

- (I^{''''}) If "N" is a proper name uttered by a speaker S at a time t, then "N" refers to x if and only if x is the one and only one object such that
- (i) S's utterance of "N" at t is causally connected with the use of the name by a group of speakers who believe and intend that they are using the name with the same reference.
 - (ii) There is a causal connection between S's utterance of "N" at t and the object x.
 - (iii) S has x in mind in uttering "N" at t.
 - (iv) S has x in mind because either the dominant intention is the intention to use the name "N" to refer to the object x which is the object referred to by the group of speakers with whose use of the name "N" S's utterance of "N" at t is causally connected; or because the dominant intention is the intention to refer to a pre-conceived object x.
 - (v) The causal connection between S's state of mind and the object x must be such that there is some characteristic which the causal chain possesses and by virtue of which x is the object at which the chain terminates.

3. TERMINATION: Before we examine the causal theory of names as it has been formulated by us in condition (I^{''''}) above, something has to be said about the way in which a causal chain terminates at a given

individual. According to Kripke, there are two ways in which the initial referential connection is established between the name and the object named. The first is by ostension and the second is by means of reference-fixing descriptions. Cases of naming where someone introduces a name for an object by ostension are more frequent than those cases where a name is introduced by means of reference-fixing descriptions. What are the conditions which have to be satisfied in order for an object to be named by ostension? One obvious condition is that in order to name an object, we must somehow succeed in naming that particular object and not some other one. Thus, when we name an object by ostension there must be some relation that obtains between us and the object named and that same relation must not obtain between us and some other object as well, during the same act of naming. For in that case we would not succeed in naming the object which we intended to name. It seems, then, that in order to identify an object for baptismal purposes a person must be in direct cognitive contact with the object. Once someone has named an object by ostension the name is passed from link to link of a chain of communication. What joins each link to the next is its causal connection with it, together with the persistent intention to use the name with the same reference as the previous speaker. At each link in the chain there is also the intention to refer to a pre-conceived object and this intention in most cases will coincide with the intention to use the name with the same reference as the previous speaker but it need not. The pre-conceived object is the object which has been named by ostension or the object with which the speaker is

acquainted. So the intention to refer to a pre-conceived object is the intention to refer to a particular object, i.e., that object which was bestowed with a name at a baptismal ceremony.

It is interesting to note that Kripke's account of naming by ostension is very similar to the account given by Russell of naming by acquaintance. Thus, Russell says:

"A name, in the narrow logical sense of the word whose meaning is a particular, can only be applied to a particular with which the speaker is acquainted, because you cannot name anything you are not acquainted with. You remember, when Adam named the beasts, they came before him one by one, and he became acquainted with them and named them. We are not acquainted with Socrates, and therefore cannot name him."¹⁸

Russell is here giving an account of what we may call "baptism by acquaintance". According to Russell we can name, i.e., give logically proper names, only to those objects of which we have knowledge by acquaintance, that is, with which we are directly acquainted. Russell's account, however, only applies to cases of ostension. But Kripke admits another way in which a name may attach to an object. The second way in which a name may attach to an object is one which we have already encountered in the second chapter. In this case an object is endowed with a name by means of a definite description which fixes the reference of the name but is not synonymous with it. Naming by means of a reference-fixing descriptions is quite different from ostension in that the person who bestows the name on the object does not have to be directly related to the object which is being given the name. Thus, for example, I can fix the reference of the name "Newman I" by means of the description "the first child to be

born in the twenty-first century". However, most cases of naming involve ostension. Thus, for example, Kripke says:

"Two things should be emphasized concerning the case of introducing a name via a description in an initial baptism. First, the description used is not synonymous with the name it introduces but rather fixes its reference ... Second, most cases of initial baptism are far from those which originally inspired the description theory. Usually a baptizer is acquainted in some sense with the object he names and is able to name it ostensively."¹⁹

We can incorporate the above account of how an object acquires a name into condition (I''''') by adding the following clause to it.

- (vi) A causal chain may be said to terminate at an object x if and only if a referential connection has been established between a name and the object x either by ostension or by means of a reference-fixing description.

The addition of clause (vi) to condition (I''''') gives us the final version of the causal theory of names.

4. SOME OBJECTIONS: In this sub-section we will consider some objections to the causal theory of names as it has been formulated by us in condition (I'''''). More precisely, we will argue that the causal theory of names does not provide a correct principle for the decipherment of names. Let us recall that the principle of decipherment is provided by thesis (v). Thesis (v) asserts that a speaker has a particular object in mind in uttering a name at a given time, if the causal connection between the speaker's state of mind and the object which the speaker has in mind is such that there is some characteristic which the causal chain possesses and by virtue of

which the object which the speaker has in mind is the object at which the chain terminates. The speaker, of course, does not have to know anything about the characteristic possessed by the causal chain by virtue of which his utterance of a name is the utterance of a name of a particular object. But we in order to decipher the speaker's use of a name must know what characteristic is possessed by the causal chain by virtue of which the speaker's use of a name is a name of that particular object. This, however, is just the difficulty. In actual practice it is highly dubious that we ever trace out in detail a chain of communication of the kind required by the causal theory of names, which we would have to do to discover some characteristic which would help us to decipher the speaker's use of a name. The notion of reference as explained in terms of a chain of communication has no bearing on the way in which we do, as a matter of practice, decipher the use of a name by a speaker. We very rarely establish with certainty, and sometimes are unable to establish even with probability the existence of such a chain of communication. In other words, we cannot trace the use of a name back to its first introduction. And, even in those cases where we can trace the use of a name back to its introduction, such etymological research usually plays no part in the way in which we decipher a sentence containing a name. So thesis (v), which was meant to provide a principle of decipherment, is false. What about thesis (ii)?

Thesis (ii) asserts that there is a causal connection between the speaker's utterance of a name and the referent of the name. Although in need of clarification as to the nature of the causal

connection, we may let this thesis stand. For the causal theorist, however, this thesis plays no part in deciphering the use of a name by a speaker. And it is with the problem of decipherment that we are primarily concerned. Thesis (ii), on the other hand, is meant to eliminate the notion of reference which task was not accomplished by thesis (i). And it does so because it asserts that a connection exists between a speaker's use of a name and the original act of baptism, by means of which a particular object acquired the name which the speaker uses. But thesis (ii) is hardly, illuminating. For thesis (ii) asserts that a speaker's use of the name "Zeno", for example, refers to Zeno just in case a causal connection exists between the speaker's utterance of the name and the object Zeno. But this does not tell us which one of the two objects which bear the name "Zeno" is the referent of the speaker's use of the name "Zeno". For that we have to turn to thesis (v) which, as we have seen, is false.

The other theses play no role in deciphering the use of a name by a speaker and they can survive. For example, thesis (vi) asserts that a causal chain may terminate at an object in either of the following two ways. It may terminate at an object if a referential connection has been established between a name and the object by ostension; or it may terminate if a referential connection has been established between a name and the object by means of a reference-fixing description. This thesis describes a viable procedure, due to Kripke, for introducing a name into our language and there is no reason for rejecting this thesis.

To sum up. We can say that the causal theory of names as it has been formulated by us is to be rejected on the grounds that it does not provide a correct principle of decipherment for referentially ambiguous names. Moreover, it is difficult to see how to modify the causal theory of names so as to get a correct principle of decipherment. Instead of attempting to modify the causal theory of names we will, in the next section, present and defend a modified version of the cluster theory of names which preserves the rigidity of names.

IV

THE AMENDED VERSION OF THE CLUSTER THEORY OF NAMES

1. KRIPKE'S COUNTER-EXAMPLES RE-EXAMINED: Before we attempt to modify the cluster theory of names as it was presented by us in Section (I), we will re-examine Kripke's counter-examples to the cluster theory of names. Let us recall that Kripke's counter-examples to the cluster theory of names were meant to show that either the uniqueness condition was not satisfied and yet the speaker could be said to refer, or the circularity condition was violated without preventing the speaker from picking out the referent of the name. Consider, for example, thesis (3) which states that if most or a weighted most of the properties are satisfied by a unique object y , then y is the referent of "N". Kripke gives an example to show that although most of the ϕ 's are in fact satisfied by a unique object, still, on the basis of our intuitions, we might judge (in some cases) that y is not the referent of "N". Thus, for example, a speaker who associated with

the name "Gödel" merely the description "the prover of the incompleteness of arithmetic" would nonetheless be referring to Gödel and saying something false of him in uttering "Gödel proved the incompleteness of arithmetic" if an unknown Viennese by the name of Schmidt had in fact constructed the proof which Gödel had subsequently stolen. Thus, according to Kripke, most people can only explain the reference of the name "Gödel" by means of the single description "the man who proved the incompleteness of arithmetic". But in spite of this, intuitively, we feel that the speaker's sentence "Gödel did not (after all) prove the incompleteness of arithmetic" would be true on our supposition.

Now, if the only means of determining the reference of the name "Gödel" is by means of the description "the man who proved the incompleteness of arithmetic" then it does not make sense to say that Gödel did not prove the incompleteness of arithmetic. For, in that case we would ask "who, then, was Gödel?" And to this question it is no good replying "the man to whom the proof of the incompleteness of arithmetic is commonly attributed". For, that would be circular as noted by Kripke himself:

"All of us in the community are trying to determine the reference by saying 'Gödel is to be the man to whom the incompleteness of arithmetic is commonly attributed? None of us will get started with any attribution unless there is some independent criterion for the reference of the name other than 'the man to whom the incompleteness of arithmetic is commonly attributed'. Otherwise all we will be saying is 'we attribute this achievement to the man to whom we attribute it', without saying who that man is, without giving any independent criterion of the reference, and so the determination will be circular."

The fact that we can entertain the possibility that Gödel did not prove the incompleteness of arithmetic without asking the question "who, then, is Gödel?" shows that there is some independent criterion for determining the reference of the name "Gödel". This would have to be true about any name about whose referent a great deal is known by at least some members of the linguistic community. So that the independent criterion for identifying the referent of the name "Gödel" would be provided by those members of the community who know a great deal about Gödel. What we do in the case of a name like "Gödel" where all that most of us know about Gödel is that he proved the incompleteness of arithmetic, is exploit the fact that the name is part of a common language. We use the name knowing that the bearer of the name could be more precisely identified by some members of the community who know a great deal more about Gödel than we do. In other words, there will be some individuals in the community who have a sufficiently rich concept of the individual Gödel (i.e., a large enough set of beliefs about him) so as to be able to single him out uniquely. The referent of the name "Gödel" will be that individual who best fits the cluster of descriptions which the most knowledgeable members of the community associate with the name "Gödel".

Similar considerations apply to the other counter-examples which Kripke gives; for example, his counter-examples to thesis (2). Thesis (2), it will be recalled, states that one or some of the properties associated with every designating expression are believed by the speaker to pick out some individual uniquely. One of Kripke's counter-examples to this thesis is the following. According to Kripke

a man on the street may want to refer to Einstein and he will do this by using the description "the man who discovered the theory of relativity". Here he has given a unique description which is supposed to pick out Einstein. But if we ask him what the theory of relativity was he might reply "it was the theory that Einstein discovered". In other words, he is very likely to violate the non-circularity condition. Kripke wants to maintain, however, that he has in this case used the name "Einstein" as a name of Einstein. That is, he has referred to Einstein. Kripke's point is that people can use a name like "Einstein" by associating the description "the man who discovered the theory of relativity" with it and yet without being able to state the theory of relativity.

It is important to note that the above counter-example to the cluster theory of names relies on the fact that if the speaker is unable to state the theory of relativity then he violates the non-circularity condition. But to ask that a person be able to state the theory of relativity is to ask for too much. All that is required is that a person be able to discover a statement of the theory and be able to identify it as the referent of the name. Moreover, the identification of the item which satisfies the description "the theory of relativity" does not have to take place via the person who first formulated the theory. Here again we are exploiting the fact that the name is being used by a community of speakers and that the referent of the name could be more precisely identified by other people in the linguistic community if necessary.

At this point we may be able to get some leverage on the

problem by appealing to the notion of a concept. Thus, when we say that someone has a sufficiently rich concept of an individual we are to be taken to mean that that person has a large enough set of beliefs about the individual in question. Since different people may make different assertions using the same name, it follows that understanding of the concept of the individual which is the referent of the name will vary from one person to another. Using this notion of a concept we can say that someone who uses the name "Einstein" by associating with it the assertion "Einstein discovered the theory of relativity" does not have to be able to state the theory before we can say that that person has grasped the concept of the referent of the name "Einstein". All that is required is that he be able to identify a statement of the theory as the referent of the name "Einstein's Theory". And this identification may proceed via the knowledge of other people in the linguistic community to which he belongs. What this means is that some people in the linguistic community have a sufficiently rich concept of the individual Einstein even though anyone speaker who uses the name "Einstein" may only have a partial grasp of that concept.

There do seem to be cases, however, where we have to say a speaker has no understanding of the concept of the individual which is the referent of the name he uses. Evans has identified these cases as ones where the "mouthpiece syndrome" (as Evans calls it) is in operation, i.e., where "we attach sense and reference to a man's remarks only because we have someone else speaking through him; as we might with a messenger, carrying a message about matters of which

he was entirely ignorant".²¹ But there does not seem to be any sharp distinction between this case and those cases where we would say that a speaker has a sufficiently rich concept of an individual whose name he uses. There seem to be a whole spectrum of cases in between, as, for example, with Kripke's scenarios regarding the name "Einstein" and "Gödel". In general it is very difficult to identify those cases where we may say that a speaker has acquired a complete grasp of the concept of an individual. How much must a speaker know before we can say that he has acquired a mastery of the semantics of the name in the sense of knowing what individual is named? Take, for example, a name like "Yevtushenko". If all that someone knows about Yevtushenko is that he is a poet, we could hardly say that he has acquired a mastery of the semantics of the name. However, it is not clear how much a speaker must know about Yevtushenko before we can say that. Is it enough if the speaker knows that Yevtushenko was a Russian and that he wrote the poem "Babi Yar"? But we all use names where we know very little about the referent of the name and yet would be surprised if we were told that we had not acquired a mastery of the semantics of the name. In all these cases we rely on the fact that the names employed by us are part of a common language.

Given the above account it is easy to see how to deal with the rest of the counter-examples that Kripke gives to the cluster theory of names. Consider, for example, thesis (5). Thesis (5) states that it is known a priori by the speaker that if x exists, x has most of the \emptyset 's. Kripke's counter-example to this thesis is the following:

"Suppose that someone says that no prophet ever was

swallowed by a big fish or a whale. Does it follow, on that basis, that Jonah did not exist? There still seems to be the question whether the Biblical account is a legendary account of no person or a legendary account built on a real person. In the latter case, it's only natural to say that, though Jonah did exist, no one did the things commonly related to him. I choose this case because while Biblical scholars generally hold that Jonah did exist, the account not only of his being swallowed by a big fish but even going to Ninevah to preach or anything else that is said in the Biblical story is assumed to be substantially false. But nevertheless there are reasons for thinking this was about a real prophet."²²

Kripke's scenario seems to go like this. To start with we use the name "Jonah" (for whatever reason) as the name of the prophet who was swallowed by a whale, who preached at Ninevah, etc. Here we are following the Biblical account. Now, we can imagine the Biblical scholars telling us that the Biblical account is a legendary account of a real person. In such circumstances, we would be led to say that Jonah did not do any of the things that he is supposed to have done on the Biblical account. Thus, Kripke wants to say that we were all the time using the name "Jonah" as the name of the person who did not do any of the things which he was said to have done in the Bible. However, if we were to apply the unamended version of the Cluster theory of names, we would be led to say that initially "Jonah" had no referent. But this runs counter to our intuitions.

To deal with Kripke's scenario we have to fall back on the notion of a name being part of a common language. As Kripke himself points out, the reason it makes sense to say that Jonah did not preach at Ninevah, ... etc. is because there are independent reasons for thinking that the Biblical account is not a pure legend about an

imaginary character but one about a real character, i.e., there is some evidence for thinking that the Biblical account of Jonah is an imaginary account of a real person. So in this case, too, we have to rely for the reference of the name "Jonah" on those speakers in the community who know much more about Jonah than we do, i.e., those speakers who possess evidence for Jonah's historicity. As an example, we can cite the author H.L. Ginsberg from whose book The Five Megilloth and Jonah, Kripke derives the evidence for saying that the Biblical account of Jonah is an imaginary account of a real person. In general, there will be some people in the linguistic community which employs the name "Jonah", who possess an independent criterion for identifying the bearer of the name.²³ For if no one possessed a criterion for identifying the referent of the name there would be no such thing as the commonly agreed referent of a name.

Now, what about thesis (6)? Thesis (6), it will be recalled, states that the statement, "If x exists, then x has most of the ϕ 's" expresses a necessary truth in the idiolect of the speaker. It is quite clear that we have to give up thesis (6). For thesis (6), as we have already seen in the second chapter, has the consequence that names are not rigid designators, that they do not necessarily designate the same object in all counterfactual situations. So thesis (6) has to be given up. It does not follow that we have to give up the cluster theory of names to preserve the lack of the non-rigid uses of a name. For one can maintain that the reference of a name is determined by applying the Principle of Charity to the corpus of utterances in which the name occurs, and at the same time maintain

that on subsequent occasions the name is used rigidly to refer to the same object in all counterfactual situations.

We are now in a position to offer an amended version of the cluster theory of names. It is important to keep in mind the fact that the modified version of the cluster theory of names which we will present is only meant to deal with the hard cases, such as those which Kripke invokes against the cluster theory of names, i.e., those cases where a speaker knows very little about the referent of his use of a name. Wilson's cluster theory of names suffices to deal with the standard cases where a speaker has a sufficiently rich concept of an object which is the referent of the speaker's use of a name. Let us, then, offer the following as a first approximation of the amended version of the cluster theory of names.

(C) If "N" is a proper name uttered by a speaker S at a time t, then "N" refers to x if and only if x is the one and only one object such that:

(i) There is a linguistic community C whose members use "N" to refer to x.

(ii) Reference in any particular case being taken to rely on the satisfaction by x of some open sentences associated with "N" by those members of C who have a sufficiently rich concept of x.

2. WEIGHTING: Clause (ii) of condition (C) above asserts that if a speaker uses a name "N" to refer to an object x at a given time t, then reference is determined by the set of open sentences which some speakers belonging to the linguistic community associate with the name

at *t*. Alternatively, given that predicates hook on to properties, we can say that whether a speaker is referring to an object with such and such a name at a given time is determined by the cluster of properties associated with the name at that time. The properties in question have to be non-question-begging properties.²⁴ The problem we are faced with is how these properties may be used to determine what speakers are referring to with names on particular occasions. As we saw in the first section, the suggestion made by Wilson is that we should apply what he calls the "Principle of Charity" or the criterion of "best fit". According to this criterion, a speaker refers to an object with a name if and only if that object satisfies more of the open sentences which the speaker associates with his use of the name than any other object. There are, however, some cases in which this criterion does not work. For example, those cases in which it is clear that a speaker is referring to different individuals on different occasions with the same name, even though the open sentences associated with the different uses of the name are the same. A good example of this is provided by Donnellan's "Aston-Martin" case which he describes in "Proper names and Identifying Descriptions".²⁵ A student meets someone at a party whom he mistakenly regards as the famous philosopher J.L. Aston-Martin. The student knows a few things about the philosopher Aston-Martin such as the fact that he is the author of Other Bodies, etc. The man whom the student meets at the party is also called "J.L. Aston-Martin". At the party the student engages the man in a conversation which has nothing to do with philosophy:

"Imagine, then, a subsequent conversation with his

friends in which the student relates what happened at the party. He might begin by saying, 'Last night I met J.L. Aston-Martin and talked to him for almost an hour'. To whom does he refer at this point? I strongly believe the answer should be 'to the famous philosopher', and not, 'to the man he met at the party'. What the student says is simply false; a friend 'in the know' would be justified in replying that he did not meet J.L. Aston-Martin, but someone who had the same name Suppose, however, that the audience contains no such doubting Thomases and that the rest of the party was of sufficient interest to generate several more stories about what went on. The student might use the name 'J.L. Aston-Martin', as it were, incidentally. For example, '... and then Robinson tripped over Aston-Martin's feet and fell flat on his face' ... In these subsequent utterances to whom was the speaker referring in using the name 'Aston-Martin'? My inclination is to say that here it was to the man he met at the party and not to the famous philosopher."²⁶

This fact suggests that a speaker may attach differing weights to the members of a cluster associated with a name on different occasions, and what is being referred to depends on these weights. Intuitively, the idea is that though the set of open sentences associated with a name may be the same on different occasions of the use of the name, the speaker may place greater or less emphasis on the satisfaction of some open sentences by the object on one occasion than he does on another, depending on what the point of his use of a name on the occasion in question is.²⁷ Thus, we need a weighting procedure before we can tell how the referents of a speaker's use of a name are to be identified. Now, the notion of a "weight" is fuzzy and it may well be, as Wilson points out, that it is not susceptible to being treated systematically.²⁸ However, it can be made more precise for a specific example. Consider the example from Donnellan quoted above. Suppose that Donnellan is right and that the first time the student uses the

name "Aston-Martin" he refers to the philosopher Aston-Martin whereas the second time he uses the name he refers to the man he met at the party. If this is so then in the two cases he is attaching different weights to the assertions. My suggestion is that this should be explained by invoking a distinction between dominant and secondary beliefs. This needs clarification.

In the case where the student utters the sentence "Last night I met J.L. Aston-Martin and talked to him for almost an hour", the student uses the name "Aston-Martin" with some belief about the man he met at the party, say, the belief that the man he talked to for almost an hour was the author of Other Bodies. But this belief can be regarded as secondary in the sense that it is seen as deriving from other beliefs about Aston-Martin which would be expressible as "Aston-Martin is the author of Other Bodies", etc. It is because the student believes that Aston-Martin is the author of Other Bodies and that Aston-Martin was the man he met at the party that he believes that he talked to Aston-Martin for almost an hour at the party. When a persons having a belief or a set of beliefs could be used to explain the possession of other beliefs by him, we will say that the former set of beliefs are dominant and the latter set of beliefs are secondary. Now, in trying to decipher the use of the name "Aston-Martin" by the student in the sentence "Last night I met J.L. Aston-Martin and talked to him for almost an hour", we can appeal to the distinction between dominant and secondary beliefs. Briefly, we adopt the hypothesis that in hard cases, the secondary beliefs do not play any role in deciphering the uses of names on given occasions. The reason

for accepting this hypothesis is that it gives us results that appear to be intuitively correct, as in Donnellan's example above.

We have to, of course, allow for the possibility that beliefs which were regarded as dominant on one occasion of a speaker's use of a name may be relegated to the status of secondary beliefs. Thus, in Donnellan's example, the student's use of the name "Aston-Martin" in sentences describing the events that occurred later on in the party, may be taken to refer to the man he met at the party. In other words, the beliefs of the student which he would express as "... and then Robinson tripped over Aston-Martin's feet and fell flat on his face", etc. are to be accorded the status of dominant beliefs. Thus, the reference of the name in this case is to the man over whose feet Robinson tripped and fell. The secondary beliefs would be those expressible as "Aston-Martin is the author of 'Other Bodies'," etc. Now, at first glance this may seem strange. How, one may ask, can the student believe that Aston-Martin is the author of "Other Bodies" because he believes that Robinson tripped over Aston-Martin's feet and fell flat on his face? Or, to put it another way, how can the belief that Robinson tripped over Aston-Martin's feet be regarded as partly explaining the belief of the student that Aston-Martin authored "Other Bodies"? However, to look at the matter in this way is already to prejudge the issue. That is, this way of looking at the matter assumes that the name "Aston-Martin" is being used by the student to refer to the philosopher Aston-Martin when he describes the subsequent events that occurred at the party. If we grant, however, that the ~~name~~ on these occasions of its use refers to the man the student meets

at the party then we can explain the apparent discrepancy in the following way. The student's subsequent beliefs about Aston-Martin (such as the belief that Robinson tripped over his feet) are to be regarded as dominant because they explain partly why the student then believes, although falsely, that Aston-Martin, i.e., the man he met at the party, authored the book Other Bodies. Similarly with other beliefs about Aston-Martin that he held before the party. It will not be possible, in most cases, to identify the precise moment at which one set of beliefs ceased to be dominant and another set took their place. But, in general, which set of beliefs is to be regarded as dominant and which secondary can be determined by context. The answer would depend upon what the point of a speaker's use of a name on the occasion in question is. We can now re-state condition (C) so as to accommodate those cases where weighting takes place.

(C') If "N" is a proper name uttered by a speaker S at a time t, then "N" refers to x if and only if x is the one and only one object such that:

- (i) There is a linguistic community C whose members use "N" to refer to x.
- (ii) Reference in any particular case being taken to rely on the satisfaction by x of some dominant open sentences associated with "N" by those members of C who have a sufficiently rich concept of x.

The above reformulation of condition (C) is to be regarded as our final version of the amended cluster theory of names. The Principle

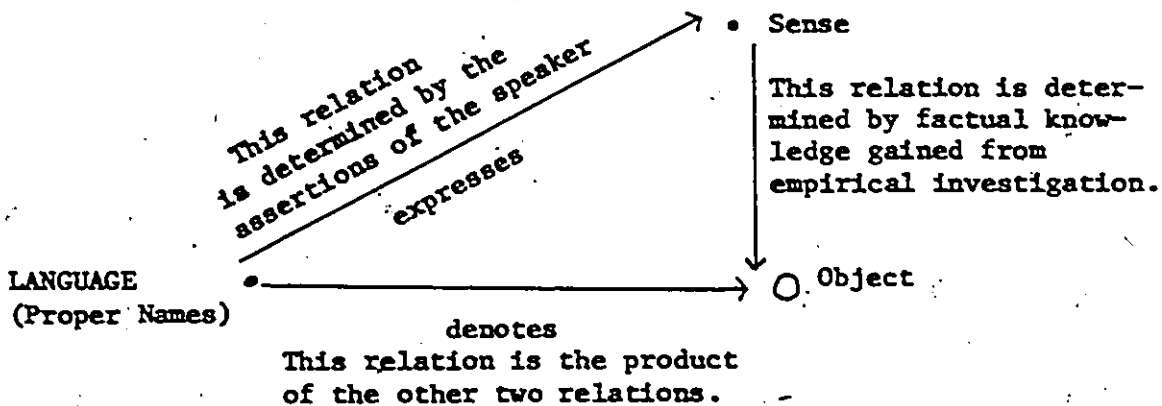
of Charity is to be applied to the set of dominant open sentences.

We can now sum up the conclusions which we have reached in this chapter. First, of all, we saw that the unamended version of the cluster theory of names is open to Kripke-type counter-examples which are meant to demonstrate that either the uniqueness condition is not satisfied and yet the speaker may be said to refer or the circularity condition is violated without preventing the speaker from picking out the referent of the name. Secondly, we have seen that the causal theory of names which was put forward by Kripke to handle his counter-examples to the cluster theory of names and to preserve the lack of non-rigid uses of names, is open to a serious difficulty. It does not provide a correct principle of decipherment for referentially ambiguous names. Finally, we presented the amended version of the cluster theory of names which preserves the lack of the non-rigid uses of names and which also accounts for Kripke's counter-examples to the unamended version of the cluster theory of names.

SUMMARY: In this thesis our main concern has been to formulate a theory of names according to which names are rigid designators, that is, a theory of names according to which the meaning of a name is not equivalent to the meaning of some particular description or even to a cluster of descriptions. However, it is important to note in this connection that Quine has been vindicated in a certain sense. For, one of Quine's objections to quantifying into modal contexts was that it commits the modal logician to essentialism. Now, although

it is not clear why this is to be regarded as an objection to modal logic, the doctrine that names are rigid designators does commit one to some form of essentialism. For, as we saw in the second chapter, the theory of rigid designators presupposes that the individual named has an individual essence (whatever it may be) and it and only it has that essence in every possible world.

If we were asked to contrast the Fregean picture of reference and the direct reference picture with our own, we could do so by means of the following diagram.



There are two ways in which the above picture of reference differs from the Fregean picture of reference. First of all, the above picture is not meant to suggest that there is such an entity as the sense of a proper name. As pointed out by Wilson the sense of a proper name may simply be taken to be the set of assertions which a speaker is willing to make in using the name. Secondly, in order to apply the Principle of Charity we have to conduct an empirical investigation to determine which object best fits the descriptions which a speaker associates with his use of a name. We don't, as the Fregean picture

suggests, look for an object which uniquely falls under a given concept.

FOOTNOTES

¹Wilson, N.L. "Substances without Substrata", The Review of Metaphysics, 12 (1959), pp. 521-539.

²Ibid., p. 531.

³Ibid., p. 533.

⁴Ibid., p. 534.

⁵Kripke, S. "Naming and Necessity", in Semantics of Natural Language, D. Davidson and G. Harman (eds.) (D. Reidel Publishing Company, Dordrecht - Holland, 1972), p. 285.

⁶Ibid., p. 262.

⁷Ibid., pp. 294-295.

⁸Ibid., p. 348, note 36. It is not clear why Kripke thinks that the name "Gödel" is not being used attributively in the sentence "Gödel proved the incompleteness of arithmetic". It seems that Kripke wants to say that in the case of a sentence like "Gödel used a diagonal argument in the proof", it can happen that the main interest of a speaker who utters the sentence is to refer to whoever proved the theorem. But then it is not clear why a speaker cannot use the name "Gödel" attributively in the sentence "Gödel proved the incompleteness of arithmetic" if he is interested in talking about whoever proved the incompleteness of arithmetic.

⁹Ibid., p. 285.

¹⁰Ibid., p. 288.

¹¹Ibid., p. 297.

¹²Ibid., p. 302

¹³Evans, G. "The Causal Theory of Names", Part I of a symposium with J.E.J. Altham, Proceedings of the Aristotelian Society, Supplementary Volume 37 (1973), p. 196.

¹⁴Kaplan, D. Demonstratives (Draft #2) mimeographed, UCLA Philosophy Department, pp. 93-99.

¹⁵Kripke, S. "Naming and Necessity" in Semantics of Natural Language, op. cit., p. 343, note 3.

¹⁶ Kaplan, D. Demonstratives, op. cit., pp. 96-97. Kaplan seems to think that we can account for the mixed intentions by either making a distinction between speaker's reference and semantic reference or by regarding one of the two intentions to be the dominant one. It did not occur to him that we would need to do both.

¹⁷ Kripke S. "Naming and Necessity" in Semantics of Natural Language, op. cit., p. 346, note 22. Kripke credits the idea that the Fregean notion of sense be replaced by that of a causal chain to Hartry Field.

¹⁸ Russell, B. Russell's Logical Atomism, ed. D. Pears (Collins Sons & Co., London, 1972), p. 56.

¹⁹ Kripke, S. "Naming and Necessity" in Semantics of Natural Language, op. cit., p. 349, n. 42.

²⁰ Ibid., p. 297.

²¹ Evans, G. "The Causal theory of names", Proceedings of the Aristotelian Society, op. cit., p. 192.

²² Kripke, S. "Naming and Necessity", Semantics of Natural Language, op. cit., p. 282.

²³ It is, of course, possible that those people in our community who possess an independent criterion for identifying Jonah derive their evidence for the historicity of Jonah from a reference to him in II Kings, as suggested by Kripke. In that case the author(s) of II Kings would have to be included in the community of speakers under consideration, since they possess a sufficiently rich concept of Jonah.

²⁴ Donnellan, K. "Proper Names and Identifying Descriptions" in Semantics of Natural Language, op. cit., p. 365. As examples of question begging properties Donnellan gives the following:

- (a) "the entity I had in mind".
- (b) "the entity I referred to".

²⁵ Ibid., pp. 356-379.

²⁶ Ibid., pp. 370-371.

²⁷ This has already been noted by Donnellan. Cf. "Proper Names and Identifying Descriptions", Semantics of Natural Language, op. cit., P. 371.

²⁸ N.L. Wilson. "Substances without Substrata", Review of Metaphysics, op. cit., p.

BIBLIOGRAPHY

- Boer, S.E. "Reference and Identifying Descriptions", The Philosophical Review, 81 (1972), 208-28.
- Carnap, R. Meaning and Necessity, Chicago, 1947.
- Church, A. Review of Quine's "Notes on Existence and Necessity", Journal of Symbolic Logic, 8 (1943), 45-7.
- "A Formulation of the Logic of Sense and Denotation", in Structure and Meaning: Essays in Honour of Henry M. Sheffer, (eds.) P. Henle et al., New York, 1951, 3-24.
- Davidson, D. "The Method of Extension and Intension" in The Philosophy of Rudolph Carnap (ed.) Schilpp, P.A., Illinois, 1963, 311-349.
- (ed., with G. Harman) Semantics of Natural Language, Dordrecht, 1972.
- Donnellan, K. "Reference and Definite Descriptions", The Philosophical Review, 75 (1966), 281-304.
- "Putting Humpty Dumpty Together Again", The Philosophical Review, 77 (1968), 203-15.
- "Proper Names and Identifying Descriptions", Synthese, 21 (1970), 335-58. Reprinted in Semantics of Natural Language, D. Davidson and G. Harman (eds.), Dordrecht, 1972, 356-79.
- "Speaking of Nothing", in Hockney, Harper and Freed (eds.), Contemporary Research in Philosophical Logic and Linguistic Semantics, Dordrecht, 1975, 93-118.
- Dummett, M. Frege: Philosophy of Language, London, 1973.
- Evans, G. "The Causal Theory of Names", Proceedings of the Aristotelian Society, Suppl. vol. 47 (1973), 187-208.
- Frege, G. "On Sense and Reference" in Philosophical Writings of Gottlob Frege, translated and edited by P.T. Geach and M. Black, Oxford, 1970, 56-78.
- Hintikka, J. "Quantifiers in Deontic Logic", Societas Scientiarum Fennica, Commentationes Humanarum Literarum 23, no. 4, Helsinki, 1957.

- Hintikka, J. "Existential Presuppositions and Existential Commitments", The Journal of Philosophy, 56 (1959), 125-37.
- "Towards a Theory of Definite Descriptions", Analysis, 19 (1959), 79-85.
- "Modality and Quantification", Theoria, 27 (1961), 119-28.
- "The Modes of Modality", Acta Philosophica Fennica, fasc. 16 (1963), 66-67.
- "Individuals, Possible Worlds and Epistemic Logic", Nous, 1 (1967), 33-62.
- Kaplan, D. "Quantifying In", in Words and Objections: Essays on the Work of W.V. Quine, (eds.) D. Davidson and J. Hintikka, Dordrecht, 1969, 178-214. Reprinted in Reference and Modality, (ed.) L. Linsky, Oxford, 1971, 112-44.
- "Bob and Carol and Ted and Alice" in (eds.) J. Hintikka, et. al., Approaches to Natural Language, Dordrecht, 1972, 490-518.
- "Trans-World Heir Lines", paper presented at a joint symposium of A.P.A. in Chicago, 1967.
- "D that" in Syntax and Semantics: Vol. 9 Pragmatics, (ed.) P. Cole, New York, 1973, 221-243.
- Demonstrations (Draft #2), mimeographed, UCLA Philosophy Department, 1977.
- Kripke, S. "A Completeness Theorem in Modal Logic", Journal of Symbolic Logic, 24 (1959), 1-15.
- "Semantical Analysis of Modal Logic" (Abstract) Journal of Symbolic Logic 24 (1959), 323-24.
- "Semantical Analysis of Modal Logic I: Normal Modal Propositional Calculi" in Zeitschrift für mathematische Logik und Grundlagen der Mathematik, 9 (1963), 67-96.
- "Semantical Considerations on Modal Logic", Acta Philosophica Fennica, 16 (1963), 83-94. Reprinted in Reference and Modality, (ed.) L. Linsky, Oxford, 1971, 63-72.
- "Identity and Necessity", in Identity and Individuation, (ed.) M. Munitz, New York, 1972.
- "Naming and Necessity" in Semantics of Natural Language, D. Davidson and G. Harman (eds.), Dordrecht, 1972, 253-355 and 763-9.

- Lewis, D. "Counterpart Theory and Quantified Modal Logic", The Journal of Philosophy 65 (1968), 113-126.
- Linsky, L. "Reference, Essentialism and Modality", The Journal of Philosophy, 66 (1969), 687-700.
- (ed.) Reference and Modality, Oxford, 1971.
- (ed.) Semantics and the Philosophy of Language, Illinois, 1972.
- Mackay, A. "Mr. Donnellan and Humpty Dumpty on Referring", The Philosophical Review, 77 (1968), 197-202.
- Quine, W.V.O. "Whitehead and the Rise of Modern Logic" in The Philosophy of Alfred North Whitehead (ed.) Schilpp, P.A. Illinois, 1941, 125-64.
- Review of Russell's Inquiry into Meaning and Truth, Journal of Symbolic Logic, 6 (1941), 29-30.
- "Notes on Existence and Necessity", The Journal of Philosophy, 40 (1943), 113-27. Reprinted in Semantics and the Philosophy of Language, (ed.) L. Linsky, Illinois, 1972, 77-95.
- "The Problem of Interpreting Modal Logic", Journal of Symbolic Logic, 12 (1947), 43-48.
- Word and Object, New York, 1960.
- "Reference and Modality" in From a Logical Point of View, New York, 1961, 139-57. Reprinted in Reference and Modality, (ed.) L. Linsky, Oxford, 1971, 17-34.
- "Three Grades of Modal Involvement" in The Ways of Paradox and Other Essays, Cambridge, 1976, 158-77.
- "Reply to Professor Marcus" in The Ways of Paradox and Other Essays, Cambridge, 1976, 177-84.
- Russell, B. "On Denoting", Mind, 14 (1905), 479-93. Reprinted in Logic and Knowledge, (ed.) R.C. Marsh, London, 1956, 48-51.
- Principles of Mathematics, New York, 1938.
- Russell's Logical Atomism, (ed.) D. Pears, London, 1972.
- Searle, J. "Proper Names", Mind, 67 (1958), 166-73.
- "The Problem of Proper Names". A section of Speech Acts: An Essay in the Philosophy of Language, Cambridge, 1969. Reprinted in Semantics, (eds.) Steinberg, D. and Jakobovits, L. Cambridge, 1971, 134-42.

- Smullyan, A. "Modality and Description", Journal of Symbolic Logic, 13 (1948), 31-7. Reprinted in Reference and Modality, (ed.) L. Linsky, Oxford, 1971, 35-43.
- Strawson, P.F. "On Referring", Mind, 67 (1950), 320-44.
- Introduction to Logical Theory, London, 1952.
- Individuals: An Essay in Descriptive Metaphysics, London, 1959.
- "Identifying Reference and Truth Values", Theoria, 30 (1964), 96-118.
- van Fraassen, B. "Singular Terms, Truth-value Gaps, and Free Logic", The Journal of Philosophy, 63 (1966), 481-95.
- Formal Semantics and Logic. New York, 1971.
- Whitehead, A. and Russell B. Principia Mathematica, London, 2nd ed., 1973.
- Wilson, N.L. "Designation and Description", The Journal of Philosophy, 50 (1953), 369-83.
- "In Defence of Proper Names Against Descriptions", Philosophical Studies, 4 (1953), 72-8.
- "Property Designation and Description", Philosophical Review, 64 (1955), 389-404.
- "Existence Assumptions and Contingent Meaningfulness", Mind, 65 (1956), 336-45.
- "Substances Without Substrata", Review of Metaphysics, 12 (1959), 521-39.
- Wittgenstein, L. Philosophical Investigations, translated by G.E.M. Anscombe, Oxford, 1971.
- Woods, J. "Essentialism, Self-Identity and Quantifying In" in Identity and Individuation, (ed.) M. Munitz, New York, 1971, 165-198.
- Vendler, Z. "Thinking of Individuals", Nous, 10 (1976), 35-46.