INFORMATION TO USERS

This was produced from a copy of a document sent to us for microfilming. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help you understand markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure you of complete continuity.

2. When an image on the film is obliterated with a round black mark it is an indication that the film inspector noticed either blurred copy because of movement during exposure, or duplicate copy. Unless we meant to delete copyrighted materials that should not have been filmed, you will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed the photographer has followed a definite method in "sectioning" the material. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For any illustrations that cannot be reproduced satisfactorily by xerography, photographic prints can be purchased at additional cost and tipped into your xerographic copy. Requests can be made to our Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases we have filmed the best available copy.

University
Microfilms
International

310 N. ZEEB ROAD, ANN ARBOR, MI 48106
18 BEDFORD ROW, LONDON WC1R 4EJ, ENGLAND
MORELLI, RALPH A.

AN EXAMINATION OF RECENT PROPOSALS IN THE THEORY OF REFERENCE

University of Hawaii Ph.D. 1979

University Microfilms International 300 N. Zeeb Road, Ann Arbor, MI 48106 18 Bedford Row, London WC1R 4EJ, England
AN EXAMINATION OF RECENT PROPOSALS
IN THE THEORY OF REFERENCE

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF THE
UNIVERSITY OF HAWAII IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN PHILOSOPHY
AUGUST 1979

By
Ralph Morelli

Dissertation Committee:
Irving M. Copi, Chairman
Chung-ying Cheng
Winfield Nagley
Edward Harter
Ty Pak
ACKNOWLEDGMENTS

This dissertation would not have been possible without the help and encouragement of my teachers, friends and family. I am especially grateful to Professors Irving Copi and Edward Harter who gave unselfishly of their time and patience to discuss and critique my work. I sincerely appreciate their many helpful suggestions and much needed encouragement.

I would also like to acknowledge and thank my parents and family for their many years of encouragement and support from afar. Finally, I shall always be indebted to Choong-lan, my wife, for sharing my frustrations during the past two years and for her constant warmth and reassurance.
ABSTRACT

This dissertation examines recent proposals addressed to a problem in the theory of reference: viz., what determines the referent of a word? Traditional answers to this question have relied on the "intensionalist principle"--i.e., the principle that the intension of a word (or the descriptive backing in the case of a proper name) determines its extension or denotation. Recently, however, Saul Kripke, Keith Donnellan and Hilary Putnam have raised objections to this principle and have proposed theories which attempt to explain reference without relying on the intensionalist principle.

The major shortcoming of traditional intensionalism is identified herein as due to conceiving the 'intension' both conceptualistically--as something which guides the speaker's use of a word--and essentialistically--as comprised of the essential attributes of the things to which the word applies. It is shown that Jerrold Katz' version of intensionalism, which is representative of the traditional approach, conceives of intensions in this way. Several problems with this approach are identified and an interpretation of 'intension' which disassociates it from both conceptualism and essentialism is suggested. The result is a view of the intensionalist principle which is immune to many of the criticisms raised by its detractors. In particular, Putnam's "Twin Earth" examples are discussed and shown to fail in their attempt to refute the intensionalist approach.

Kripke's "rigid designation" thesis is viewed as an alternative to Russell's "disguised description" theory, and both Kripke's and Putnam's
defenses of rigid designation are discussed. It is maintained that, like Russell's theory, rigid designation represents an interesting recommendation for how we might interpret the designators of formal modal logics but fails as a theory of how proper names (and other designators) are used in ordinary discourse. Putnam's defense of rigid designation, which is seen to conflict with several points of his "sociolinguistic hypothesis"—Putnam's alternative to the intensionalist thesis, is shown to rely on a faulty notion of the "extension" of a word. Kripke's arguments in support of rigid designation are shown to depend on our intuitions about the contingency or necessity of certain English sentences, and, since our intuitions may vary, his arguments are judged inconclusive. A similar objection is raised against Kripke's belief in the existence of contingent a priori and necessary a posteriori statements.

Kripke's "causal theory" and Donnellan's "historical explanation theory" are both attempts to provide an alternative to the Fregean "sense theory" of proper names. It is maintained herein that the key determinant of the referent of the use of a proper name is the speaker's intention to refer to a unique individual. Because they are based on a rejection of the thesis that the referent is determined by the "descriptive backing" of a name, Kripke's and Donnellan's theories give a less-than-central role to "speaker's intention" and are found wanting on that basis. It is concluded that the recent proposals by Kripke, Donnellan and Putnam do not provide a superior basis (to intensionalism) for an understanding of the connection between words and their referents.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>iii</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>iv</td>
</tr>
<tr>
<td>CHAPTER I INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>CHAPTER II TRADITIONAL INTENSIONALISM</td>
<td>12</td>
</tr>
<tr>
<td>The Classical View and Its Problems</td>
<td>12</td>
</tr>
<tr>
<td>Katz' Theory of Meaning</td>
<td>24</td>
</tr>
<tr>
<td>CHAPTER III PUTNAM'S VIEWS ON MEANING</td>
<td>41</td>
</tr>
<tr>
<td>Putnam's Attack on Intensionalism</td>
<td>41</td>
</tr>
<tr>
<td>Putnam's Account of Rigid Designation</td>
<td>53</td>
</tr>
<tr>
<td>CHAPTER IV THE INTENSIONALIST APPROACH TO PROPER NAMES</td>
<td>80</td>
</tr>
<tr>
<td>Fregean Sense and Reference</td>
<td>80</td>
</tr>
<tr>
<td>Mill's Theory of Proper Names</td>
<td>85</td>
</tr>
<tr>
<td>The Backing of Descriptions</td>
<td>91</td>
</tr>
<tr>
<td>Kripke's Interpretation of the Principle of Identifying Descriptions</td>
<td>95</td>
</tr>
<tr>
<td>CHAPTER V RIGID DESIGNATION</td>
<td>112</td>
</tr>
<tr>
<td>Rigid Designation and Formal Semantics</td>
<td>112</td>
</tr>
<tr>
<td>Rigid Designation as a Theory of Ordinary Proper Names</td>
<td>118</td>
</tr>
<tr>
<td>Response to an Objection</td>
<td>127</td>
</tr>
<tr>
<td>CHAPTER VI CONTINGENCY, NECESSITY AND RIGID DESIGNATION</td>
<td>138</td>
</tr>
<tr>
<td>The Contingent A Priori?</td>
<td>138</td>
</tr>
<tr>
<td>The Necessary A Posteriori?</td>
<td>147</td>
</tr>
<tr>
<td>CHAPTER VII THE ATTACK ON THE PRINCIPLE OF IDENTIFYING DEScriptions</td>
<td>160</td>
</tr>
<tr>
<td>Donnellan's Argument</td>
<td>160</td>
</tr>
<tr>
<td>Some Purported Counterexamples</td>
<td>175</td>
</tr>
<tr>
<td>CHAPTER VIII</td>
<td>THE HISTORICAL AND CAUSAL THEORIES OF REFERENCE</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td></td>
<td>An Exposition of Donnellan's and Kripke's</td>
</tr>
<tr>
<td></td>
<td>Theories</td>
</tr>
<tr>
<td></td>
<td>Genuine Alternatives to the Intensionalist</td>
</tr>
<tr>
<td></td>
<td>Approach?</td>
</tr>
<tr>
<td></td>
<td>Some Problems with Donnellan's and Kripke's</td>
</tr>
<tr>
<td></td>
<td>Proposals</td>
</tr>
<tr>
<td>CHAPTER IX</td>
<td>CONCLUDING REMARKS</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER I
INTRODUCTION

It is obvious that, in some sense, language is 'about' the world: we use verbal and written symbols to refer to things in the world, to attribute properties to those things, to report relations between them, and so on. What is not so obvious, and what has, indeed, generated some controversy recently, is the precise relationship between the expressions of our language and the world. An important question presents itself: what sort of relationship must obtain between a referring expression and its referent in order that the expression may be used to refer to the referent?

Traditional answers to this question have relied on the principle that intension determines extension. They have held that associated with any significant expression is some meaning or concept or property or, more generally, an intension, by virtue of which that term denotes a unique individual or class of individuals. For instance, Frege argues that there is connected with every sign "the sense of the sign, wherein the mode of presentation is contained."¹ One example that Frege uses is that the sense of 'Aristotle' might be the pupil of Plato and teacher of Alexander the Great.² Basically, Frege's view is that the sense of the name 'Aristotle' serves as a criterion of identity for the referent of the name. Other proponents of the Fregean approach to proper names—for example, Searle and Strawson—have viewed the sense of proper names in terms of a "backing of descriptions" which serves, again, to pick out or identify the referent of the name.³
Similarly, with regard to general terms, it has been maintained, traditionally, that an expression deserves to be applied to certain individuals only if those individuals display the attributes which make up that term's intension. For example, Mill argued that the word 'man' connotes the attributes of corporeity, rationality, animal life and a certain human form, and he argued that nothing deserves to be called 'man' unless it possesses those attributes. Again, according to the traditional view, the extension of a term, in this case those individuals to which the term may correctly be applied, is determined intensionally--i.e., by checking each individual to see that it possesses the attributes associated with the term.

Recently, however, the traditional approach to reference has been challenged. With respect to proper names, for example, Donnellan has recently argued that the principle of identifying descriptions--the principle that the referent of a name is picked out or identified by its descriptive backing--is false. He offers several examples which purport to show that it is possible to refer to a unique individual even if no significant number of the descriptions associated with that individual's name are true of that individual. According to Donnellan, successful reference depends not upon a backing of descriptions, but upon what he calls a "historical connection" between the name and its referent. Similarly, Kripke has argued that the referent of a proper name is determined not by a descriptive backing but by a causal link between the name and its referent. In either case, the test for the referent is an extensional link between the name and its referent.
Closely related to this view of proper names is the position taken by Kripke and Putnam with respect to the meaning of natural kind terms --i.e., terms that purport to denote things that occur naturally in the sense that they are not human artifacts. It is argued that a natural kind term is similar to a proper name in that it is connected directly to its extension. In their view, the extension of a natural kind term is not determined, as Mill held, by its intension or connotation. Rather, the extension is determined by the essence of the things to which it is applied. As Schwartz puts it, "[w]hat determines whether some stuff is gold is its atomic structure."\(^6\)

Kripke has invented the notion of a "rigid designator" which he defines as a term that denotes the same object(s) in all possible worlds. What this means is that rather than being determined by the associated properties or the descriptive backing, which may pick out different individuals in different possible worlds, the extension of a rigid designator is determined by the essence or the internal structure (atomic or genetic) of its individual members. An example often cited is that the term 'water' has as its extension all and only those wholes with the chemical structure H\(_2\)O. According to proponents of the rigid designation theory, the term 'water' has always designated the same stuff, even through prior to the work of Priestly and others, it was not known that water is H\(_2\)O.

We have here, then, two conflicting views about the connection between language and the world. We shall refer to the traditional approach, both as it applies to proper names and general terms, as the intensionalist view since it is based on the principle that the referent
is picked out by the intension. We shall refer to the new approach as the 
extensionalist view since it is based on a rejection of the intensionalist 
principle.

It is the purpose of this dissertation to sort out the issues 
involved in this dispute and thereby provide an indication of which is the 
more adequate theory of reference. My thesis is that the recent proposals 
by Kripke, Donnellan and Putnam are very problematic. First, the argu­ments presented by proponents of the extensionalist approach against the 
traditional view are misleading. I shall attempt to show that very 
often these arguments are directed against inaccurate and improperly 
developed portrayals of the intensionalist view. I conclude that these 
arguments do not support the view that the intensionalist theory should 
be abandoned. Second, I shall attempt to point out what appear to be 
serious difficulties with the arguments marshalled to support Kripke's, 
Donnellan's and Putnam's positive theories. In short, I believe that 
the traditional approach to reference, with its reliance on intensional 
meaning properly developed, is essentially sound and much superior to the 
more recent proposals.

This is not to say that traditional intensionalism is without faults 
of its own. In Chapter II, I identify its major shortcoming as the 
conflation of two distinct philosophical doctrines: conceptualism and 
essentialism. In the course of the intensionalist tradition, these two 
doctrines have become intertwined with the notion of intensional meaning, 
with the result that intensions have come to be viewed both as mental in 
nature and role, and as comprised of concepts of the essential attributes 
of things. In part, the idea that meanings are in the mind has been at
the bottom of the recent rejection of the intensionalist approach, particularly as regards Putnam's arguments. However, I argue that intensionalism can be disassociated from both conceptualism and essentialism; the result is a version of intensionalism that provides a quite defensible interpretation of the principle that intension determines extension.

In Chapter III, I discuss Putnam's attempted refutation of the principle that intension determines extension and argue that his attack fails to refute that principle. I employ the distinction between subjective and conventional intension to show that Putnam's arguments are directed against a straw man—viz., the principle that subjective intension determines extension. I also point out a shortcoming in Putnam's positive theory of meaning. In the second part of this chapter I take up Putnam's defense of the rigid designation theory and argue that it is incompatible with his sociolinguistic hypothesis (which represents Putnam's alternative to the intensionalist thesis). A key contention of my argument is that Putnam's view is based on a faulty and misleading notion of the extension of a word.

Chapter IV is an exposition of the intensionalist approach to proper names and concentrates on providing a clear explanation of the principle that names require a backing of identifying descriptions. I discuss Frege's arguments for his sense theory of names as well as Russell's "disguised description" theory of names and contrast these with Mill's view on proper names. Mill's view represents an anticipation of Donnellan's and Kripke's approach to proper names, and it is important, therefore, to explore its shortcomings. Also, I discuss Kripke's
interpretation of the traditional approach, arguing, contra Kripke, that one must distinguish between Russell's disguised description theory, which might reasonably be called a theory of meaning for proper names, and Frege's sense theory, which is clearly a theory of reference. I show that Searle's theory follows in the Fregean tradition and is immune, therefore, to the criticisms raised by Kripke against Russell's view interpreted as a theory of meaning.

Chapters V and VI constitute a critique of Kripke's rigid designation theory and his corollary belief in the existence of contingent a priori and necessary a posteriori truths. My central claim, briefly, is that rigid designation is, at best, a useful recommendation for formal modal logics but fails as a theory about ordinary proper names.

In Chapter VII, I discuss Kripke's and Donnellan's criticisms of the principle of identifying descriptions--i.e., the principle that the backing of descriptions is what picks out the referent of a proper name. I view reference within the speech-act framework and argue that Donnellan's arguments, in particular, ignore the role which the speaker's intentions play in determining the referent of the use of a proper name. I show that Donnellan's arguments lead to the patently absurd view that a speaker can successfully refer to an individual to whom he did not intend to refer and about whom nothing is known. Finally, in Chapter VIII, I attempt to show why Kripke's and Donnellan's own positive accounts are mistaken.

Before proceeding, it will be useful to touch briefly on the positive insight that emerges from the criticism of Putnam's, Kripke's and Donnellan's proposals: a proper account of the referential aspect of
language must be based on the principle that intension determines extension. In certain respects, this is tantamount to reaffirming the validity of the intensionalist position, but it must be noted that through the course of the criticism of traditional intensionalism, a revised version of intensionalism is developed. The central insight of this revised intensionalism is the disassociation of the notion of intension from the sway of conceptualism and essentialism. This is brought about by defining intensions as clusters of attributes that may be invoked to pick out or describe the extensions of words. According to this interpretation, intensions are viewed neither as the mental guides of speech behavior, nor are they assumed to be comprised of the so-called essential attributes. By thus disassociating intensions from mentalism and from essentialism, we set them free to play a pedagogical role, which is consistent with the fact that native speakers can explain and justify their choice of words, but cannot describe, in psychological terms, what takes place when they employ those words.

Regarding my reliance on the notion of intension, a caveat is in order. I have not attempted to answer Quine's claim that we lack precise identity conditions for intensions. Quite simply, Quine's worry is not my worry in this dissertation. Quine's attack on intensions is motivated largely by a problem that arises in the philosophy of logic and set theory: we possess rigorous identity conditions for sets (the formal counterparts of extensions) but not for functions (the formal counterparts of intensions). For this reason, primarily, Quine banishes intensions from his philosophy of language and attempts to get along with extensions. Although I have some doubts about the success of Quine's venture, I shall
not attempt to air them in this dissertation. To do so would take us too far afield. The justification for this is simply that the opponents of intensionalism discussed herein direct their attacks at the intensionalist principle and not at the notion of intension per se. In particular, they do not appear to share Quine's ontological scruples as regards intensions.

This is not to say that Quine's worry has absolutely no bearing on the discussion of the referential aspect of natural language. In the wake of Quine's attack on intensions it is now generally agreed that intensions are not as easy to characterize as formerly believed. But the reason for this may be that most intensions are, by their very nature, open-textured rather than fixed and definite. This suggests that we would do better to look towards Wittgenstein's "family resemblance" theory as a basis for the explication of intensions than towards more traditional proposals. 8

One final point is worth noting at the outset. As I interpret the notions of intension and extension they are correlative; one is explained in terms of the other. This is as it should be so long as we hold to the principle that intension determines extension. One of the problems with the recent proposals is that they tend to downplay the importance of intensional meaning. Regarding proper names, the recent proposals recommend dropping the "descriptive backing" in favor of historical or causal chains that link names directly to their referents. With regard to general terms, it is recommended that intensions (here in the sense of known or believed attributes of things) be dropped in favor of the essences of things (whether known or not). In either case, intensions are given short shrift.
The problem with these recommendations is that once we banish intensions we give up all hope of characterizing the abstractive, inductive nature of our symbolism. To see this we might caricature extensionalism by viewing its explication of meaning as a museum display, while intensionalism would propose a dictionary with definitions. For example, assume that the museum display for the word 'cat' is a stuffed cat. In and of itself, this display cannot provide an adequate account of the meaning of 'cat' because it does not effectively rule out the possibility that the word 'cat' refers to the particular cat in the display rather than to objects similar to that display in certain respects. Nor would the display suffice if it were expanded to include several specimens. The problem with the display is that it fails to account for the abstractive nature of symbolism. The word 'cat' conveys the meaning it does because speakers are able to abstract general forms or schemata from the particulars to which the word applies. And it is for the characterization of this abstractive process that intensions prove useful: they purport to describe, in general terms, those features common to all cats and thereby provide us with a characterization of the extension of 'cat'. A full grasp of the meaning of 'cat' involves the ability to apply the word correctly to a potential infinitude of similar objects, and this ability requires the knowledge not only that certain things have been called 'cat' but also that all things similar to those in certain respects would be called 'cat'. The museum display fails to convey this knowledge and must, therefore, be supplemented with descriptions of the extension; these of course must rely on intensions. On the other hand, dictionaries appear to serve us quite well in this regard; they do
convey the knowledge that all objects of such and such a description are called 'cats', and therefore provide a basis for understanding the abstract nature of symbolism. Thus, as we shall see, rather than deserving a diminished role in the explanation of reference (as the recent proposals suggest), intensions are quite essential for an understanding of the connection between words and their referents.
NOTES--CHAPTER I


2. Ibid., p. 58.


7. The distinction between theory of meaning and theory of reference is due to W. V. O. Quine. See his From a Logical Point of View (2nd ed., rev.; New York: Harper and Row, 1961), p. 130: "The main concepts in the theory of meaning, apart from meaning itself, are synonymy (or sameness of meaning), significance (or possession of meaning), and analyticity (or truth by virtue of meaning). Another is entailment, or analyticity of the conditional. The main concepts in the theory of reference are naming, truth, denotation (or truth-of), and extension. Another is the notion of values of variables."

8. See below, p. 23.
CHAPTER II

TRADITIONAL INTENSIONALISM

The Classical View and its Problems

The classic statement of the intensionalist theory of reference is provided by Mill's doctrine of connotation and denotation. According to Mill's theory, most general terms possess two modes of meaning which contemporary logicians call intensional meaning and extensional meaning. Mill distinguishes between connotative and non-connotative terms. The latter group includes both proper names (which are claimed by Mill to lack connotation and are said to denote certain individuals directly) and singular abstract terms (such as 'whiteness' which likewise lack connotation and merely designate or denote objects). All other general terms, according to Mill, are connotative which means that they denote a certain class of individuals and connote or imply an attribute or set of attributes:

The word man, for example, denotes Peter, Jane, John, and an indefinite number of other individuals, of whom, taken as a class, it is the name. But it is applied to them, because they possess, and to signify that they possess, certain attributes. These seem to be, corporeity, animal life, rationality, and a certain external form, which for distinction we call the human.

It is also worth noting that Mill views connotation or intensional meaning, as the meaning of a word properly so called, arguing that "whenever words have properly any meaning, the meaning resides not in what they denote, but in what they connote." According to Mill, connotative terms "convey . . . information" about the objects they denote. The information conveyed or connoted
serves as a criterion of application for the word; that is, it determines the objects to which the word may correctly be applied. This relation between the intension and the extension of a word, constitutes the basis of what we have called the intensionalist principle. Namely, the meaning or intension of a word determines its extension. Mill illustrates this principle as follows:

"Every existing thing which possessed all [the attributes connoted by 'man'] would be called a man, and anything which possessed none of them, or only one, or two, or even three of them without the fourth, would not be so called." 8

Thus, in Mill's view, the connotation of the word 'man' fixes its extension in the sense that for something to be called 'man' it is necessary and sufficient that it possess the attributes included in the connotation of 'man'. 9

Because the connotation of a word is thus connected with the necessary and sufficient conditions for its application, it is not surprising to find Mill equating the connotation with the notion of essence or essential attributes. Thus, Mill argues that "[t]he essence of man, simply means the whole of the attributes connoted by the word." 10

In thus equating meaning and essence Mill is presupposing Locke's distinction between real and nominal essence; that is, he is equating meaning and nominal essence, not meaning and real essence. Locke had previously argued that we must distinguish the essence of an object, which in many cases is unknown, from those attributes that figure in our application of the word. 11 According to Locke, the real essence of a thing is "its very being ... whereby it is what it is." 12 For material substances real essences consist of "the real internal, but generally (in substances)
unknown constitution of things, whereon their discoverable qualities depend."13 Also, because real essences of things consist of the "constitu­tion of their insensible parts" these cannot influence our applica­tion of general terms.14 Therefore our use of a general term depends on its nominal essence which, for Locke, is "that abstract idea which the general . . . name stands for."15

Locke's reasoning on this point is essentially sound in the sense that our application of words to things must depend on those qualities of things which are readily observable;16 in a similar spirit we might also say that our employment of language must depend on our beliefs about the world. For the most part, the abstract ideas which constitute the meaning of words according to Locke consist of concepts of the so-called 'superficial' attributes of things.17 For example, the word 'gold' stands for our ideas of yellow color, fusibility, malleability, heavy weight and so forth. We see then that whereas Mill conceives of meaning in terms of abstract entities--viz., attributes, Locke conceives of meaning in terms of our concepts of these attributes. Nevertheless, the difference between their respective views is minimal, and aside from their differences over the nature of meaning, both theories are intensionalist in that they are based on the principle that intension determines extension.18

In one form or another, the conceptualism or mentalism that characterizes Locke's view of meaning has been a constant feature of the intensionalist approach. It can be seen in Aristotle's claim that "[s]poken words are the symbols of mental experience" where 'mental experience' is further explained in terms of "the individual concepts,
the true universals which come to be stabilized . . . within the soul."\textsuperscript{19} The influence of this Aristotelian conceptualism can be detected in Lewis' view that each general term is connected with a "criterion-in-mind" that functions as "a schema [or] rule and an imagined result of which will determine applicability of the expression in question."\textsuperscript{20} Similarly, this Aristotelian conceptualism can be detected in Katz' still more recent pronouncement that "the meaning of a sentence can be taken as represented by some specific set of recursive rules [that represent] the mental principles (or brain mechanisms) inside the heads of speakers."\textsuperscript{21}

In one sense, the identification of meanings with concepts or criteria-of-application possesses an element of truth and is therefore quite harmless. The element of truth is that which is sometimes expressed by noting that language-use is rule-guided behavior; similarly, the learning of language is largely the learning of certain rules and conventions. However, the sense in which we can truly say that language users are guided by rules or conventions requires a word of explanation. Certainly language users are not guided by linguistic rules in the same way that one may be guided by a road map and compass or by a railroad timetable. Generally speaking, when employing language, we do not go through the kind of checking and reasoning that occurs when we use a road map; so we are not guided by linguistic rules in any conscious way. Rather, we must suppose that the rules have been internalized (and become "second nature" so to speak) during the process of acquiring one's language. Nevertheless, the fact that we are guided by rules is evidenced by our ability to justify our choice of words, by our ability
to explain the usage of words to others, and by several other such capabilities. Thus, we can safely say that language use is rule-guided behavior—and hence conceptualistic in this rather minimal sense—meaning by this that language users are able to invoke rules to justify their employment of words.

Conceptualism, in the sense just explicated, does have its limitations and must not be allowed to overstep its bounds. For one thing, because the learning of one's language is a very idiosyncratic process, we cannot expect different speakers to be guided by identical rules and concepts, even when they display virtually the same language behavior. As Lewis has noted, "it is not evident that use of the same terms on the same occasions indicates that the same sense meanings are in mind," nor, he goes on, is there any way to verify that the same sense meanings are held in common by a community of speakers.

Despite the wide circulation of familiar caveats like Lewis's, most intensionalists have not fully appreciated how limited this harmless conceptualism must be. Katz, for example, appears to allow it to get out of hand and the conclusions he draws about our ability to discover what guides a speaker's behavior are fundamentally misguided. As we shall see, he follows Locke and other intensionalists in viewing meanings as mental (or neural), and yet insists that one can obtain empirical confirmation that such and such is the meaning of a word (in this mentalistic sense). The fundamental misconception here, to employ a distinction employed by Quine in a similar connection, is the idea that we can discover what guides speech behavior. Quine argues that at best the linguist can formulate rules that fit the behavior of speakers and
may indeed test these empirically. The assumption that such rules also 
guide their behavior is illegitimate and unconfirmable in principle.\textsuperscript{24} 
The reason for this is that two different rules may generate precisely 
the same behavior just as two different mathematical functions may 
generate precisely the same solution sets for a given domain and range. 
I shall develop this point more fully below, but for now suffice it to 
say that Quine's insight is not inconsistent with the idea that language-use 
is rule-guided. The point of his argument is a methodological one: 
we possess no empirical methods for discovering precisely what guides 
our use of words.

Although the above version of conceptualism is relatively harmless, 
it does contribute to problems when combined with a second philosophical 
doctrine that has come to be associated with the intensionalist approach: 
Platonic (and later, Aristotelian) essentialism. This is the view 
expressed recently by Baylis, that "[t]he existence of communicable 
knowledge requires shared meanings which in its simplest form, is 
knowledge of the common characters exhibited by various objects and 
events."\textsuperscript{25} Throughout the intensionalist tradition this knowledge of 
'common characters' has been characterized in various ways, but the same 
basic idea can be found in the thought of most philosophers who have 
adressed what we would now call the problem of meaning.

Prior to Locke, the attributes that go to make up what Mill calls 
the 'connotation' of a general term, were viewed as constituting the 
(real) essence or essential attributes of things. As we know, the notion 
of essence originated with Plato, and it played an important role in the 
search for answers to the so-called 'Socratic questions'—i.e., such
questions as 'What is virtue?' and 'What is justice?' With regard to how one should answer such questions, Socrates' advice to Meno has had an important influence on all subsequent inquiries into meaning: "Even if [the virtues] are many and various, yet at least they all have some common character which makes them virtues. That is what ought to be kept in view by anyone who answers the question, 'What is virtue?'" This 'common character' of which Socrates speaks is what Aristotle later defined as the essence of a thing. Similarly, it is what Locke was later to call the real essence of things.

It is important to note that the Socratic questions were not intended, primarily, as inquiries into the meaning of words; for the most part, Socrates assumed that we already know the meaning of words like 'justice' in the sense that we know how to apply them correctly to actions or situations being evaluated. Rather, the Socratic questions were invitations to characterize or describe virtue itself. These 'invitations to characterize' are sometimes described as being invitations to define things, not words. As A. E. Taylor points out, "[f]rom the Greek point of view, the problem of definition itself is not one of names, but of things." Thus, unlike contemporary views of definition, definition for the Greeks was an investigation into the very nature of things, not primarily, an investigation into the use of words.

Although the Greeks did not distinguish sharply between meaning and essence, Locke did so distinguish between them, at least insofar as their roles are concerned. If, following Copi, we distinguish between theoretical definition and lexical definition, we can see that the Greeks were interested in the former, while Locke's notion of nominal
essence more closely resembles the latter. As Copi defines them, theoretical definitions are attempts to formulate a scientifically adequate characterization of the objects to which a given word is applied, while lexical definitions are pedagogical in nature and are intended to report the usage of words. Locke's notion of nominal essence, being directed as it is towards the explanation of word usage and not towards the description of the natures of things, clearly resembles the notion of lexical definition. However, unlike lexical definitions, which because they play a pedagogical role need not be based on essential attributes, Locke's nominal essences (and Mill's connotations as well) are still comprised of the so-called essential attributes. Thus, although Locke clearly differentiates the roles played by nominal and real essences, he fails to divorce his notion of lexical meaning from that of essence: meanings or intensions (or nominal essences, as Locke called them) are invoked to describe word usage, but they remain wedded to essences. In short, as Quine points out, "the Aristotelian notion of essence was the forerunner . . . of the modern notion of intension . . . Meaning is what essence becomes when it is divorced from the object and wedded to the word."  

We see then that not only has intensionalism been associated with a (rather harmless) version of conceptualism, it has also been connected with essentialism. And, as we noted above, problems do arise when these two distinct philosophical doctrines become mixed. We might describe the problem that results as follows. Conceptualism is usually invoked to account for a language user's linguistic competence: a speaker regularly applies a word correctly because he has learned some rule or other that
guides his application. The limitation involved in this view is that the
learning and application of rules is a highly idiosyncratic process. How
then on this model do we account for the fact that language users under-
stand each other? It is at this point that some form of essentialism
creeps into the picture. As we saw with Baylis' remark, communication has
often been accounted for by appealing to "shared meanings." For
example, the following remarks by Locke have set the tone for many sub-
sequent accounts of communication:

The chief end of language in communication being to be
understood, words serve not well for that end, neither in
civil nor philosophical discourse, when any word does not
excite in the hearer the same idea which it stands for in
the mind of the speaker.

Although there is no explicit mention of essences in Locke's remarks, the
idea that knowledge of nominal essences must be shared by several speakers
in order for communication to take place is implicit in Locke's view; we
need only remind ourselves of Locke's reliance on definition--or, for
Locke, the spelling out of the nominal essence--as the remedy for the
abuses and imperfections of language. Thus, the upshot of the conflation
of conceptualism and essentialism is Locke's view that communication
depends (ideally at least) on the sharing of nominal essences. In short,
we see that the notion of an intension (or nominal essence) has been
pressed into two different kinds of service: to explain individual
linguistic competence and to explain communication. The appropriate con-
clusion appears to be that the notion must be relieved of at least part of
this burden.

We might view the problem that results from this conflation as a
tension between a basically subjective, idiosyncratic phenomenon--a
speaker's linguistic competence—and a basically intersubjective phenomenon—communication. Traditionally, the distinction between subjective and conventional connotation has been relied on to resolve this tension, and I shall do likewise. However, it will be necessary to define 'conventional connotation' in a way that disassociates it from both conceptualism and essentialism.

The distinction between subjective and conventional connotation has been variously drawn. But I shall adopt the following. The subjective connotation of a word for a speaker consists of those attributes believed by that speaker to be possessed by the objects to which he would apply the word. This definition preserves that harmless sense of conceptualism while allowing for the idiosyncratic nature of an individual's linguistic competence. The conventional connotation, as I shall employ this notion, is not connected necessarily with any particular speaker's beliefs (conceptualism) nor, necessarily, with any particular set of attributes (essentialism). Rather, we shall define it as a cluster of attributes which could be invoked to describe the application of a word to a particular extension, where the extension of a word is defined as those objects "to which the word may correctly be applied." The reason I adopt this definition should be transparent: it avoids both conceptualism and essentialism. Few problems arise for the intensionalist so long as intensions are viewed operationally; that is, as concepts used to describe the linguistic conventions of a given language community. (By 'conventions' here I mean simply that certain words are applied correctly only to certain objects and not others.)
In addition, if we conceive of intensions in the way suggested, we can give a clear sense to the intensionalist principle: native speakers can (ideally at least) invoke a cluster of attributes to describe how a certain word should be applied. This is a large improvement on the idea that some concept guides the speaker's use of a word for the simple reason that it does not require us to worry about what is, in principle, hidden from view. In an analogous but slightly different form, we can also state the intensionalist principle as follows: a linguist may describe the conventions of a given language by invoking clusters of attributes that pick out the objects to which the words apply.

When viewed in the way suggested, the conventional connotation is little more than a notion designed to characterize the lexical meaning of words. I am suggesting then that it be viewed as that which is expressed by a lexical definition, whose primary purpose, as we saw, is to describe word usage. Quite obviously, when so viewed, it makes little sense to talk about the conventional connotation, since various sets of attributes may be invoked to pick out a given word's extension. To take a traditional example, both 'rational animal' and 'featherless biped' pick out equally well—or, perhaps we should say 'equally badly'—the extension of 'human'. In preferring one of these over the other, the linguist may be guided by such (non-empirical) considerations as relative simplicity or relative pedagogical superiority, but cannot be guided by the fact that one definition is more empirically correct than the other; so long as the linguist's aim is to report the usage of words, no such standard of empirical correctness is available.

Two points are worth noting regarding the above definitions. First, the cluster of attributes that go to make up the conventional connotation
must be viewed as open-textured. That is, it must not be assumed that a given cluster of attributes will be common and peculiar to all things to which a given word is applied. To take Wittgenstein's familiar example, not all games possess some particular cluster of attributes by virtue of which we call them 'games'. Rather, among games we find what he calls "family resemblances," which are "a complicated network of similarities overlapping and criss-crossing."38

Second, it would seem from the above definitions that the claim that intension determines extension is a mere truism; for, as we have defined it, the connotation serves to pick out or determine the extension. As I suggested in the Introduction, this is as it should be; for, to describe adequately the abstractive nature of our use of language, we need both intension and extension. In that sense, these notions are correlative. There is a problem here, however; for it is doubtful that those, like Putnam, who reject the intensionalist principle would quarrel with this truism. It must be, therefore, that they do not take the force of the principle as we have just characterized it. The problem, as I see it, resides with a difference in the respective definitions of 'extension'. As here defined the extension of a word is tied to its intension; it is those objects which satisfy the intension. As we shall see, however, Putnam characterizes the notion of extension somewhat differently. In his view the extension is those objects to which the word has been applied. We shall deal in some detail with Putnam's views below, but for now it will suffice to indicate briefly what I take to be the inadequacy of his notion of extension. The problem is that the extension of a word does not just include those objects to which the word
has, in fact, been applied; it must include those objects to which the word would be applied. Only in this way can it be a measure of the linguistic competence of language users which includes, of course, the ability to apply words correctly to objects they have not yet encountered. We shall return to this point below.

To conclude, the chief merit of the proposed explication of 'intension' is that it enables us to avoid the problems engendered by the mixing of conceptualism and essentialism without having to abandon the principle that intension determines extension.

Katz' Theory of Meaning

Katz' theory of meaning represents the most recent attempt to construe meaning along the conceptualistic-essentialistic model. According to Katz, the empirical linguist should not be satisfied with mere lexical definition, as we have suggested. He points out, correctly it would seem, that lexical definitions cannot provide the basis for a theory of meaning in the traditional sense, since they do not give us the meanings of words:

There are several reasons why the semantic component of a grammar must employ theoretical definition [in Katz' sense of this notion] rather than lexical definition. One is that defining words by pairing them with synonymous expressions in the same language (or even a different one, it makes no difference here) cannot provide an account of their meaning. Such pairings say no more than that the meaning of this or that word is the same as the meaning of some expression without at all saying what the meaning of either is . . . Accordingly, lexical definitions cannot give the meaning of their definienda.

It is evident from this passage that Katz does believe that one can discover the meaning of the expressions of our languages. His hopes rest on what he calls "theoretical definitions." As Katz explains,
theoretical definitions define terms on the basis of constructs from linguistic theory, that theory in empirical linguistics which expresses what is common to natural languages in the form of a definition of the notion 'natural language'. That is, a term from a natural language is defined by a dictionary entry which represents each of its senses in the form of a theoretical construction, called a 'reading', that is composed, not of words from that language, but of constructs, called 'semantic markers', drawn from the theoretical vocabulary of empirical linguistics.42

As Katz explains in another place, the theoretical constructs--the semantic markers--that form the basis of his theory of meaning are frankly mentalistic. They express "a speaker's inner, private thoughts or ideas."43 Similarly, his semantic markers--Katz' analogue for intensions--represent "the class of equivalent ideas that we . . . have in mind . . ."44 In addition to its reliance on mental entities as the basis for meaning, Katz' theory is also essentialistic, being based on the assumption that we can distinguish the purely semantic information that governs our use of words from the rest of our beliefs about the world.

Several authors have pointed out that except for couching his theory in empirical-sounding technical jargon, Katz has gone but little way beyond traditional, conceptualistic theories of meaning. For example, Wilson claims that "Katz has got himself involved in a mythological psychology, which would be described (anathematized) as conceptualism."45 Therefore, in this section, we shall be concerned primarily with the essentialistic side of Katz' theory, our aim being to show that it enjoys no more empirical justification than its traditional forbears.

According to Katz, the problem of determining the meaning of a word boils down to the problem of distinguishing the purely semantic information that governs our use of a word from the rest of our beliefs
about the world.\textsuperscript{46} It is precisely this distinction between the semantic and non-semantic (or empirical) that Quine's picture of language rules out. According to Quine's view, no empirical sense can be made "of a distinction between what goes into a [native speaker]'s learning to apply an expression and what goes into his learning supplementary matters about the objects concerned."\textsuperscript{47} The distinction between semantic and non-semantic information is parallel to the distinction between the analytic and synthetic. For, if it is held that, say, 'female parent' constitutes the purely semantic information involved in our use of 'mother', it would follow that 'Mothers are female' is analytic. On the other hand, if 'once having had a mother' is regarded as non-semantic information, then 'All mothers once had a mother' would be synthetic.\textsuperscript{48} It is not surprising, therefore, that Katz' attempt to provide a basis for his distinction between semantic and non-semantic information is tied up with his attempt to provide a firm basis for the analytic-synthetic distinction. It will be useful, therefore, to look briefly at Quine's attack on this distinction.

To begin with, Quine does not deny that the truth of some sentences is more dependent than others on experience. For example, he would admit that in the face of an anomalous experience we would be less likely to revise our belief that cats are animals than our belief that it is now raining in Manoa; the latter belief is more closely tied to experience than is the former. Quine admits that truth "depends on both language and extralinguistic fact" but denies that the truth of an individual sentence can be analyzed into a linguistic component and a factual component.\textsuperscript{49} According to Quine, the beliefs that provide the underpinnings
of our language do not confront reality piecemeal but as a whole; some of these are closer to the experiential frontier than others, in the sense that they are more easily abandoned or revised than others. But no belief, according to Quine, is immune to revision, including such supposedly necessary truths as 'Cats are animals'.

Quine's argument against the analytic-synthetic dichotomy sets out to show that the whole affair involves us in something like a vicious circle: To provide a boundary between the analytic and synthetic, one must isolate the meanings of the words involved; but in the process of attempting to isolate meanings, one finds that one must depend on such notions as synonymy or linguistic entailment or necessary truth which themselves depend on our ability to isolate meanings. Katz attempts anew to escape from Quine's circle, and it is to his argument that we now turn.

According to Katz, to avoid Quine's circle, one must show that the meaning, or "dictionary entries for the lexical items [of a sentence]" are correct "independently of settling the particular question of [that sentence's] analyticity." That is, in order to establish that a sentence is analytic—say, the sentence 'Mothers are female'—one must have independent grounds for saying that 'female' is part of the meaning of 'mother'. One must be able to determine the meaning of 'mother' without somewhere along the line appealing to the analyticity of 'Mothers are female' or the synonymy of 'mother' and 'female parent'. Moreover, in order to avoid the charge of vacuity, one must somehow base the definition on empirical grounds and not merely stipulate that 'mother' means 'female parent'.
According to Katz, the dictionary entries of the lexical items of natural languages can be discovered empirically. He argues that what particular entry we assign to a word depends on the particular semantic theory we adopt. So far, at least, we can agree with Katz. Moreover, he argues that native speakers are able to detect such "semantic" properties and relations between words as analyticity, synonymy, semantic ambiguity, redundancy, and so on, and it is on the basis of having observed such semantic relations that the linguist is led to assign a particular dictionary entry to a lexical item. Therefore, according to Katz, one can determine whether a piece of information belongs to the dictionary entry for a word by observing the linguistic behavior, and ferreting out the linguistic intuitions of native speakers. His criterion, then, is this: The information belongs to the dictionary entry—i.e., is part of the meaning of the item—if and only if some semantic relation—e.g., analytic entailment—cannot be accounted for without appeal to this information. What Katz intends here comes out much more clearly through example, so let's look at how he would attempt to verify the hypothesis that 'female' is part of the meaning of the word 'mother'.

Katz argues that "if the inference from (1) to (2),

(1) Xantippe was a mother.
(2) Xantippe was a female.

is accepted as valid and if its validity cannot be explained without one lexical reading of 'mother' containing semantic markers that represent the information 'is female,'" then 'female' is part of the dictionary entry for 'mother'. And Katz argues that one can determine that the above inference is valid by observing the behavior of native speakers and
by ferreting out their intuitions through questions like 'Is the inference from (1) to (2) valid and why?'

The problem with Katz' argument is that one needn't appeal to meanings at all to account for the speaker's intuition of its validity. That Katz' argument begs the question can be seen by considering the following inference:

(3) Spectacular Bid is a horse.
(4) Spectacular Bid can run faster than a turtle.

If we suppose that most speakers of English regard this inference as valid, we might be tempted, if we follow Katz' reasoning, to view 'the ability to run faster than a turtle' as part of the meaning of 'horse'. Surely, most speakers of English would believe--true the generalization 'Horses run faster than turtles' and it is this suppressed belief (or premise) that may be invoked to explain the (supposed) validity of this inference. 57 But saying this much tells us nothing about the meaning of 'horse,' because a person's belief in the validity of this inference counts as evidence only of that person's beliefs, not that person's beliefs-about-meanings. Moreover, such inferences could be admissible evidence for a speaker's beliefs-about-meanings only if we assume that they are analytic entailments. But to assume this is question-begging since we are presumed to know nothing as yet of the analyticity of 'Mothers are female'. Therefore, Katz' attempt to escape Quine's circle fails.

In response to my argument, it might be argued that sentences like 'All cats meow' and 'All cats are animals' just are not on a par. Surely we can imagine cats that have lost their ability to meow, but if we are to believe Katz, it would be impossible for there to be a non-animal
cat. Therefore, truths like 'Cats are animals' must be necessary.

The supposed impossibility of non-animal cats follows directly from the claim that analytic sentences are necessarily true. Putnam has responded to this sort of argument by noting that "[e]ven if cats turn out to be robots remotely controlled from Mars we still call them 'cats'; . . . not only will we still call them 'cats', they are cats . . .". Actually, Putnam's claim is too strong. If we should discover that cats are robots we would not be committed to calling them 'cats'. We might, for example, prefer not to abandon our belief that cats are animals and insist, instead, that there just aren't any cats. We could then invent a new word to refer to robot-cats. But this has nothing to do with the meaning of 'cat' regardless of what we decide, for we could still employ the term 'cat' to refer to those same objects which we formerly believed to be animals. The point is that the conventions which govern our use of 'cat' do not dictate a course of action for the situation described. At present, our use of 'cat' is shaped in part by the belief that those things to which the word is applied are animals, but there is no reason that this belief rather than another--e.g., 'Cats go meow'--should be the decisive factor in determining our future application of 'cat'. Moreover, with regard to the notion that the intension is a criterion of application, it would be the case, perhaps, that as we presently use the word 'cat' our conventions can be accounted for by adopting the criterion 'cats are animals of such and such outward form, such and such behavior, and so forth'. This criterion would succeed in picking out the correct extension of 'cat' as we presently
use it. There are other criteria which could have been employed to do the same job, of course. If our belief that cats are animals should come to be abandoned, however, and if we decide to let this affect our use of 'cat', this would require that the linguist adopt a different criterion to account for our now revised speech habits vis-à-vis the word 'cat'.

The claim that it would be impossible to have a cat that is not an animal has an air of plausibility about it which is due, perhaps, more to the outrageousness of Putnam-like examples than to any cogency in the claim itself. We believe perhaps that nothing could shake our belief in the fact that cats are animals. But maybe the following remarks can help dispel the air of plausibility. There is a society in England, or so I'm told, that believes the earth to be flat. They have devised a consistent conceptual scheme around this belief and can account for such apparent anomalies as the pictures of the spherical earth sent back by astronauts. Now in most contexts this society employs the word 'earth' in much the same way as the rest of us. They believe the earth is one of the planets which orbit the sun and so forth. Moreover, when they use the word 'earth' they are referring to the very same thing as the rest of us--viz., the planet we inhabit. What's more, we understand them as referring to that planet when they speak about the earth. The point is that the word 'earth' refers to a particular object that both we and the members of this bizarre society can identify by means which are independent of the question of whether the earth is round or flat. What particular properties this object has are debatable, and there is room for disagreement over even the most obvious properties. Nevertheless,
there is an overlap in our respective beliefs about the earth which enables us to understand each other as talking about the same object. The beliefs upon which we agree suffice to pick out the denotation of 'earth' and they enable us to anchor our quite disparate talk about earth to the appropriate object. Thus, so long as there is some common ground among members of a given language community, we can all talk about the same object and still entertain widely disparate beliefs about that object. We might say, then, that the common ground on which communication is based is ultimately the objects that our words denote and our beliefs about those objects but not the way each of us connects the words to the objects via our beliefs about them. However, this 'common ground' needn't be construed mentalistically. Rather, it may be viewed on the Fregean model, whereby meaning is viewed as a function of both the sense and reference of the word.

Much the same sort of situation exists with regard to our beliefs about cats. Cats are objects which we can distinguish from other objects of the world in many different ways. We have many beliefs about these objects and some of these beliefs serve to connect the word 'cat' to these objects. All of these beliefs were discovered through ordinary empirical procedures but such procedures never provide knowledge of necessary facts. Therefore, if it should be discovered that cats are robots, this would imply that sentences like 'Cats are animals' are false, but it would not imply that such sentences are not about (or do not refer to) those objects which we formerly believed to be animals. The question then of whether there could be cats which are non-animals is moot: We have things that we call 'cats' and those (ex hypothesi) could be non-animals; and that's the end of the matter.
It follows, then, that no sharp boundary exists between sentences which are necessarily true in virtue of their meaning and those which are empirically true in virtue of the testimony of experience. In the light of an anomalous experience we could abandon any of our beliefs about the world, and in that sense no one of these beliefs is any more necessary than any other. As Quine would say, so much the worse for analyticity. Similarly, and for much the same sort of reasons, no sharp boundary exists between purely semantic information (which, according to Katz, guides our use of words) and collateral knowledge about the world. Our use of language is influenced—we may justifiably say "guided"—by all our beliefs, and it is hopeless to attempt to divide these beliefs into semantic and non-semantic components.

We see then that Katz' version of intensionalism inspires no more confidence than its traditional forbears, and for familiar reasons: intensions insofar as they are seen as guides for linguistic behavior—i.e., insofar as they are construed mentalistically—cannot be construed essentialistically—i.e., as being comprised of purely semantic information. For a durable version of intensionalism, we should keep the conceptualism clearly tied to the individual language users (via subjective intensions) and we should drop the essentialism altogether; (conventional) intensions are best viewed as clusters of attributes and can be quite useful when so viewed, but the clusters need not be essences.
NOTES--CHAPTER II


2. Throughout this thesis I shall employ the expressions 'intension' and 'connotation' and 'extension' and 'denotation' synonymously. Throughout the intensionalist tradition these expressions have been defined in various ways by different logicians, and in some cases 'intension' and 'connotation' and 'extension' and 'denotation' were not used synonymously. However, contemporary logicians appear to employ them interchangeably. For example, see Irving M. Copi, *Introduction to Logic* (5th ed.; New York: Macmillan Publishing Co., Inc., 1978), pp. 142-146. For a tabulation of the different characterizations that have been given to the above expressions see Joseph C. Frisch, *Extension and Comprehension in Modern Logic* (New York: Philosophical Library, 1969).


4. Ibid.

5. Ibid. I shall discuss more fully what Mill means by words implying or connoting attributes when I come to discuss his views on proper names. See below, pp. 85-90.

6. Ibid.

7. The idea that the intension or connotation serves as a criterion of application has been a traditional feature of the intensionalist approach to meaning. Thus we find Locke arguing that "[b]etween the nominal essence [Locke's notion for Mill's connotation] and the name there is so near a connexion, that the name of any sort of things cannot be attributed to any particular being but what has this essence." (John Locke, *An Essay Concerning Human Understanding*, ed. A. C. Fraser (two volumes; 1st ed.; New York: Dover Publications, Inc., 1959), III, iii, 16.) More recently, we find C. I. Lewis referring to what Mill called the connotation as a "criterion-in-mind." (C. I. Lewis, *An Analysis of Knowledge and Valuation* (La Salle, Illinois: The Open Court Publishing Company, 1946), p. 37.)


9. As Hilary Putnam has pointed out, the most obvious difficulty with this sort of claim is that "a natural kind [like man] may have abnormal members." (Putnam, 'Is Semantics Possible?' in Stephen P. Schwartz (ed.), *Naming, Necessity and Natural Kinds* (Ithaca:
Strictly speaking, then, Mill's claim must be qualified to allow for abnormal members of humankind who would not possess the attributes contained in the connotation of 'man.'


12. Ibid., III, iii, 15.

13. Ibid.

14. Ibid., III, iii, 17. According to Locke, we cannot rely on the real essences of things in ranking them into sorts or in subsuming them under a single name because "we know them not. Our faculties carry us no further towards the knowledge and distinction of substances, than a collection of those sensible ideas which we observe in them." (III, vi, 9.)

15. Ibid., III, iii, 15.

16. It would be interesting to speculate on what Locke would say if he knew, as we now know, that the "insensible parts" of things are discoverable. Today we not only know that gold things are yellow colored, malleable, fusible, etc. but also that their matter is the element with atomic number 79 (and all that that number tells us about the atomic structure of gold). One possibility open to Locke would be to view the atomic structure of gold—or our concept of it, to follow Locke's view—as part of the nominal essence of 'gold' since we now know it. But would this be in keeping with Locke's view that the nominal essence affects our use of the word 'gold'? The question is moot and would seem to depend for its answer on whether we view the scientific or technical use of the word as constitutive of its wider meaning. Locke assumed that such technical usage does not contribute to the meaning of our words, arguing that "languages, in all countries, have been established long before sciences," and that "for the most part, [words] in all languages, received their birth and signification from ignorant and illiterate people, who sorted and denominated things by those sensible qualities they found in them." (III, vi, 25.) However, Locke's view may be too unyielding, since examples abound where word usage has been affected by scientific definitions. For example, whales are only rarely (and incorrectly) referred to as fish. On the other hand, there are examples of word usage that have not been touched by scientific discoveries; for example, we still speak of the sun rising and setting.

17. The scare quotes here are intended to indicate that I regard the notion of superficial qualities as suspect. This notion is sometimes used to indicate that certain qualities are readily observable
and hence require little scientific investigation for their discovery. But this is arbitrary in view of the fact that what constituted scientific knowledge in one age has a way of becoming the popular wisdom of later generations. Also, many attributes that are often regarded as 'readily observable' are actually quite theory-laden. The assumption that a sharp boundary can be drawn between theory and observation has been vigorously criticized by Paul Feyerabend, among others, and his view has much to recommend it. See his Against Method (Atlantic Highlands, N.J.: Humanities Press, 1975).

18. Putnam goes even further and suggests that a theory like Mill's would also be mentalistic since the grasping of (abstract) connotations is still "an individual psychological act." (Hilary Putnam, 'Meaning and Reference,' The Journal of Philosophy, 70 (1973), p. 699.)

19. Aristotle, De Interpretatione, 16a-3, E. M. Edghill, trans. and Posterior Analytics, 100a-100b5, Edghill, trans.


22. As regards the connections between our concepts, words and the things of the world, Quine's words are worth bearing in mind: "Beneath the uniformity that unites us in communication there is a chaotic personal diversity of connections and, for each of us, the connections continue to evolve." (Word and Object (Cambridge, Mass.: The M.I.T. Press, 1960), p. 13.) I agree with Quine that the process by which we acquire our language is a matter of learning both the connections between words and things and the connection between new words and other words of our language. Since we all learn our language in slightly different ways, it is gratuitous to assume that we can detect those concepts (or brain mechanisms or intensions) that connect an individual word to its extension. Likewise, it is wrongheaded to assume that the intensional meaning of a particular word would be the same for all speakers. These points will loom large in our subsequent discussion.

23. Lewis, p. 143. See also Rudolf Carnap, 'Meaning and Synonymy in Natural Languages,' in Carnap, Meaning and Necessity (2nd ed.; Chicago: The University of Chicago Press, 1956), pp. 236-237, where Carnap argues that the assignment of an intension to a word is an "empirical hypothesis" that cannot be completely verified. Similarly, Katz too views the meaning of a word as a "theoretical construct." (J. J. Katz and Jerry A. Fodor, 'The Structure of a


27. See Aristotle, Metaphysics, 129b14, W. D. Ross translation.


29. For an overview of different approaches to definition see Richard Robinson, Definition (Oxford: Oxford University Press, 1954).

30. Copi, p. 140. As Copi points out in 'Essence and Accident,' (The Journal of Philosophy, 51 (1954), pp. 706-719), we can view the scientific search for theoretical definitions as a search for (real) essences in the Lockean sense (where the essence is viewed as the most fundamental properties of things in terms of which their other properties are explained--Locke, III, iii, 17). On the other hand, we can see that lexical definitions have little use for essences. Since they are primarily pedagogical, the definiens of a lexical definition must employ words (or concepts) that are familiar to those for whom the definition is intended; these will rarely be the so-called essential attributes of the thing. At best, theoretical definitions and hence essences are applicable to theory of meaning only insofar as we are interested in scientific terms or the scientific usage of everyday terms.

31. Quine, Logical Point of View, p. 22.

32. See above, p. 17.

33. Locke, III, x, 4.

34. Ibid., III, ix-III, x.

35. For an account of the various ways in which this distinction has been drawn see Frisch, op. cit.


37. It may be useful to compare this interpretation of the intensionalist theory with that offered by Stephen Schwartz who appears
sympathetic to recent extensionalist accounts of meaning. Schwartz argues as follows: "... the descriptions associated with a term do not determine what it is to be of that kind. What determines whether some stuff is gold is its atomic structure. Likewise water is H2O--some stuff is water only if it has the right chemical structure. Biological kinds are determined by genetic structure, and other natural kinds are similarly determined." (Schwartz, op. cit., pp. 27-8.) Obviously, Schwartz has a different notion of 'determines' in mind than that involved in the interpretation of the intensionalist principle offered in the text. We might characterize the difference as that between something being gold (Schwartz') and something being called 'gold' (the interpretation suggested in the text). Clearly insofar as the connotation functions as a criterion for how a word is applied by a certain community of speakers, it does not purport to be a criterion for what it is to be such and such. I shall develop this point more fully when I come to discuss Kripke's rigid designation thesis. See below, Chapter V, notes 35 and 39.


39. See below, pp. 45-46.


41. Katz' notion of 'theoretical definition' must be distinguished from Copi's which was explained on pages 18-19 above.


44. Ibid., p. 177.


46. For example, Katz ('Logic and Language,' p. 84), regards as a "false dogma" Quine's thesis that "semantic facts about words are indistinguishable from empirical facts about their referents" and proposes a criterion for distinguishing "changes of meaning and changes of belief about the world." That no such distinction between meanings and beliefs is possible has been argued by Keith Donnellan. He points out that if we distinguish 'whale1' whose meaning includes 'mammalian' and 'whale2' whose meaning does not, then upon discovery that many of the things we've been calling 'whales' do not have mammalian characteristics, we'd be at a loss to explain whether this is a change in meaning or a change in
belief: "... the supposed reason for shifting of 'whale' [to 'whale2'] is not different in kind from reasons for shifting a belief." ('Necessity and Criteria,' The Journal of Philosophy, 17 (1962), p. 656.) I believe that we must agree with Donnellan and Quine on this point and I shall present additional arguments in support of this belief presently.

47. Quine, *Word and Object*, p. 38.

48. The examples are Katz' in 'Logic and Language,' pp. 84ff.

49. Quine, *Logical Point of View*, p. 36.

50. Ibid., pp. 42ff.

51. Consider, for example, C. I. Lewis' argument on behalf of analyticity. According to Lewis, 'being a cat' entails 'being an animal' where entailment or strict implication is defined in terms of material implication and 'necessity' as follows: $p \rightarrow q = df. \Box(p \rightarrow q)$. We can then symbolize "being a cat" entails "being an animal" as follows: $(x)(x \text{ is a cat} \rightarrow x \text{ is an animal}).$ This last formula may be interpreted as 'Necessarily all cats are animals.' (For Lewis' argument see An Analysis of Knowledge ... , pp. 155-157.) In response to such an argument, Quine has pointed out that to suppose the adverb 'necessarily' makes sense is to suppose that we have already made sense of the notion of analyticity because to say 'Necessarily cats are animals' is to say no more than "'Cats are animals' is analytic." (Logical Point . . . , p. 30) In other words, to appeal to the entailment of 'being an animal' by 'being a cat' as evidence for the analyticity of 'Cats are animals' begs the question. By means of similar arguments Quine shows that appeals to synonymy or meanings also fail to provide an empirically sound, non-circular basis for analyticity.


53. In its narrowest sense a semantic theory can be regarded as an empirical hypothesis about the meaning of a word, which is then tested by appeal to the linguistic intuitions and behavior of the native speaker.


55. Ibid., p. 85.

56. Ibid.

57. For a point similar to this in spirit see Quine, *Word and Object*, p. 38, where Quine argues that "[T]he distinction between [meaning
and belief] is illusory ... It is simply a question whether to call the transitivity short cuts changes of meaning or condensations of proof, and in fact an unreal question."


59. Putnam used to agree with this and probably still would. He argues: "Once we found out that cats were created from the beginning by Martians, that they are not self directed, that they are automata, and so on, then it is clear that we have a problem of how to speak." ('It Ain't Necessarily So,' *Journal of Philosophy*, 59 (1962), pp. 660-661.)

60. This point is similar to Donnellan's point regarding a change in our beliefs about those things we call 'whales'. See note 46 above. For Katz' response to Putnam and Donnellan see his 'Logic and Language', pp. 88-96.

61. This last point is obviously in need of elaboration. See the discussion of Kripke's view that certain scientific truths are necessary *a posteriori* on pages 146-155 below, especially p. 148.
CHAPTER III
PUTNAM'S VIEWS ON MEANING

Putnam's Attack on Intensionalism

Having looked at the shortcomings of traditional intensionalism, let us now turn to the shortcomings of a theory that has arisen in response to intensionalism: Putnam's theory of meaning. Putnam's theory is difficult to classify in terms of the intensionalist/extensionalist dichotomy for it appears to adopt features of each, resulting in a view that appears inconsistent and contradictory in certain respects. We can agree with Putnam's assessment that

the traditional [intensionalist] theory has taken an account that is correct for 'one-criterion' concepts (i.e., for such concepts as 'bachelor' and 'vixen'), and made it a general account of the meaning of general names.

Yet it is not easy to understand what Putnam means by the following claim:

Meaning does not determine extension, in the sense that given the [intensional] meaning and a list of all the 'properties' of a thing (in any particular sense of 'property') one can read off whether the thing is a lemon or acid or whatever.

As we shall see, in the sort of possible-worlds framework that Putnam sets for himself, we shall be able to distinguish between something being X and something being called X. The proper role of intensional meaning, we can then say, is as an indicator of whether something would be called X by members of a certain language community, not (necessarily) whether it is X. Thus, meaning determines whether the thing would be called a lemon, etc., not whether it is a lemon. With this sort of
emendation to Putnam's above claim, one finds little reason to object to the traditional intensionalist principle. But this is to anticipate what lies ahead. Let's begin with a look at Putnam's assault on the principle that meaning (intension) determines extension.

Putnam's argument takes the form of several purported counter-examples that are designed to refute what he regards as two traditional assumptions:

(1) That knowing the meaning of a term is just a matter of being in a certain psychological state . . .

(2) That the meaning of a term (in the sense of 'intension') determines its extension (in the sense that sameness of intension entails sameness of extension). 5

Putnam's argument sets out to show that the notion of intensional meaning does not jointly satisfy these two assumptions.

According to Putnam, understanding a word, in the sense of knowing its intension, is a matter of being in a certain psychological state, because "grasping these [intensions is] still an individual psychological act."6 The conclusions we reached in the preceding chapter suggest that it is wrongheaded to view 'knowing the meaning' on the model of grasping the intension of a word: words simply do not have intensions in this singular sense. Nevertheless, there is a sense in which knowing the meaning of a word may be viewed as 'being in a certain psychological state'.

Assume that 'knowing the meaning of a word' amounts to being able to apply the word correctly. This ability presupposes the ability to detect mistaken applications of the word, which presupposes, in turn, the use of some criterion or other as a measure for correct usage. Let us call this criterion the speaker's subjective intension. Here we make
no assumption that all speakers who apply the word correctly are guided by the same subjective intension, for different criteria—e.g., 'featherless biped' and 'rational animal'—can lead to (more or less) identical applications of the same word. On the other hand, if the subjective intensions are, in fact, different, we can sometimes expect some differences in word usage among speakers. For example, a speaker who stuck rigidly to the 'featherless biped' criterion might apply the word 'human' to Diogenes' plucked chicken. In that case we might say his subjective extension—i.e., those objects to which that speaker would apply the word—differs from that of the speaker who employs 'rational animal' as criterion. The usefulness of this notion of subjective extension for differentiating the idiosyncratic speech habits of speakers will become more evident as we proceed.

The above remarks make reasonably clear the sense in which it is correct to say that 'understanding a word' or 'knowing how to apply a word' implies 'being in a certain psychological state'. The psychological state, of course, is an ability or capacity. With this in mind, let us look at Putnam's purported counter-examples to the intensionalist principle. As we shall see, all of Putnam's examples are directed at the straw-man principle that subjective intension determines extension.

Putnam's first example employs the fiction of a Twin Earth and runs as follows. Assume that there is a planet that is exactly like Earth in all respects, including the fact that people on Twin Earth speak English. It is supposed even that every speaker on Earth has a "Doppelganger" on Twin Earth who is the exact duplicate "in appearance, feelings, thoughts, interior monologues, etc." as some speaker on
Let us call two such creatures Oscar₁ and Oscar₂, Oscar₁ being an inhabitant of Earth and Oscar₂ his Twin Earth counterpart. Now we're asked to suppose that everything on the two planets is exactly alike except that what people on Twin Earth call 'water' has the chemical formula XYZ instead of H₂O (and thereby the corresponding differences in atomic structure). In all other respects XYZ is indistinguishable from H₂O at normal temperatures and pressures. Further, "the oceans and lakes and seas of Twin Earth contain XYZ and not water, ... it rains XYZ on Twin Earth and not water, etc." According to Putnam, "there is no problem about the extension of the term 'water'" in this fictitious example:

the word simply has two different meanings ... The extension of waterₘ [the sense in which it is used on Earth] is the set of all wholes consisting of H₂O molecules, or something like that; the extension of water in the sense of waterₐ is the set of all wholes consisting of XYZ molecules. Moreover, Putnam argues that in 1750, before it was discovered that waterₘ is H₂O and waterₐ is XYZ, the extensions of 'waterₘ' and 'waterₐ' were "just as much" H₂O and XYZ respectively. Because of this alleged difference in extension, Putnam concludes that Oscar₁ and Oscar₂ understood the term 'water' differently in 1750 although they were in the same psychological state, and although, given the state of science at the time, it would have taken their scientific communities about fifty years to discover that they understood the term 'water' differently. Thus the extension of the term 'water' (and, in fact, its 'meaning' in the intuitive preanalytical usage of that term) is not a function of the psychological state of the speaker by itself.

In response to this sort of example, it should be noted that the extension of a word is a measure of its use by a community of speakers, not a measure of its use by one particular speaker. At best, the
subjective intension of a particular speaker (Oscar₁ say) determines how that speaker would use the word, not how an entire language community would use the word. Thus Putnam is entirely correct in noting that subjective intension—i.e., the intension associated with 'water' by Oscar₁ and Oscar₂—does not determine the extension of 'water'. However, it is not clear that any traditionalist ever held that subjective intension does determine extension. Therefore, if we distinguish subjective from conventional intension, we can see that, strictly speaking, Putnam's example has little force against the traditional principle.

It might be argued that I have misinterpreted Putnam's example by construing the intension associated with 'water' by Oscar₁ and Oscar₂ as the subjective intension. If we take Oscar₁ and Oscar₂ as ideal speakers who know their language perfectly, then we can view the intension that they associate with 'water' as the conventional intension. On this interpretation the traditional principle appears to fail since Oscar₁ and Oscar₂ are in the same psychological state, but the respective extensions of 'waterₑ' and 'waterₐₑ' are different (according to Putnam).

One problem with this conclusion, however, is that it rests on a faulty—in fact, an untenable—notion of extension. Insofar as the notion of extension is useful for theory of meaning, it must include those objects to which a term would be applied; it is only in this way that it can be viewed as a measure of the linguistic competence of native speakers. Putnam's notion, however, construes the extension as consisting only of those objects to which the term has been applied and thus can account only for the linguistic performance of the native speaker. He then argues that intension does not determine extension because the intensions
of 'water' on Earth and Twin Earth respectively are the same but the words have been applied to different objects on each planet.

The mistake involved in this line of reasoning can be seen if we assume—as Putnam himself does—that inhabitants of Earth visit Twin Earth and vice versa. In that case, Oscar₁ and Oscar₂ would, in 1750, have applied 'water' to precisely the same things since they had not yet detected any difference between XYZ and H₂O. In short, there would have been no difference in the extension of 'water' as employed by Oscar₁ and Oscar₂ in 1750. (The extension, of course, would have been the disjunction of waterₑ and waterₚₑ.) And ex hypothesi, since the conventional intensions of 'waterₑ' and 'waterₚₑ' were exactly the same, it is true to say that in 1750 the intension of 'water' determined its extension in the sense that sameness of intension determines sameness of extension.

The correct extension of 'water' in 1750—i.e., the extension as it was used by Oscar₁ and Oscar₂—would have included all wholes having atomic structure XYZ or H₂O. It is not the case, as Putnam supposes, that Oscar₁ and Oscar₂ understood 'water' differently in 1750; their understanding of the term and the substance(s) to which it referred was exactly the same. On the other hand, once it was discovered that the stuff called 'water' on Earth has a different structure than the stuff called 'water' on Twin Earth, then assuming the isolation of the two language communities, there would be differences in the way Oscar₁ and Oscar₂ understand 'water'. But these differences in the (conventional) intensions would be accompanied by differences in extensions only if the speakers of the respective worlds decided not to apply the term 'water'
indiscriminately to the stuff in the lakes and rivers of both Earth and Twin Earth. That is, if Oscar₁ would not apply 'water' to the stuff in the lakes of Twin Earth because it doesn't have the structure H₂O, then his use of 'water' would differ from the Twin Earthian use of the term. Differences in the understanding of a term must show up in this way in the usage of the term.

Putnam's second example is equally misleading. He assumes that on Twin Earth the pots and pans, which are called 'aluminum', are really made of molybdenum so that 'aluminum₉' is used to refer to molybdenum and 'molybdenum₉' is used to refer to what Earthians regard as aluminum. He also assumes that a metallurgist from either planet could easily tell that 'aluminum₉' was used to refer to molybdenum and that 'molybdenum₉' was used to refer to aluminum. Because of this the confusion of aluminum with what is called 'aluminum₉' involves only part of the linguistic community. He then argues that typical speakers, like Oscar₁ and Oscar₂, not being metallurgists, have the same psychological states, but that 'aluminum' has the extension aluminum in the idiolect of Oscar₁ and the extension molybdenum in the idiolect of Oscar₂: "Also we have to say that Oscar₁ and Oscar₂ mean different things by 'aluminum'; that 'aluminum' has a different meaning on Earth than it does on Twin Earth, etc."₁⁶

Again, however, Putnam's example does not suffice as a counterexample to the intensionalist principle. If Putnam is trying to show that Oscar₁'s intension is the same as Oscar₂'s--i.e., that their subjective intensions are the same--but that the extension of the term is different, he is attacking a straw man. For no traditionalist, to
my knowledge, has maintained that subjective intension determines (conventional) extension. In this example it is true (by supposition) that the respective subjective intensions of 'aluminum' for typical speakers Oscar₁ and Oscar₂ are identical; but it would follow from this that their respective subjective extensions are likewise identical. That is, assuming they visit each other's planet, they would apply 'aluminum' to precisely the same objects, so long as they remained unaware of the relevant differences between aluminum and molybdenum. It is only for the respective metallurgists of Earth and Twin Earth—again by supposition—that the terms 'aluminum' and 'molybdenum' have different subjective intensions; these would correspond to differences in the way they would apply these words, however. But none of this goes against the traditional assumption that intension determines extension.

Moreover, insofar as the linguist is interested in the correct use of these words, his attempt to formulate the conventional intension of 'aluminum' on both Earth and Twin Earth would include the information that is known to the metallurgists as well as the information that is known to the typical speaker. Assume, for example, that it is agreed that the correct use of 'aluminum' excludes its application to the pots and pans of Twin Earth. This means that typical Earthian and Twin Earthian speakers alike apply it wrongly. The linguist can describe the correct use of the term by including the criterion known to the metallurgist in the conventional intension of 'aluminum'. This would pick out the correct extension of that word. And he can account for the misuse of the term by various Oscars by excluding this criterion from his characterization of their subjective intensions. In this way the
linguist's description of the conventions upon which the use of 'aluminum' is based conforms to the traditional assumption that intension determines extension. We see, then, that Putnam's example fails because of his failure to distinguish between the subjective intension of a word, which may well lead to an incorrect usage of the word but which determines that speaker's application nonetheless, and the conventional intension of the word, which is that sense of intension in which it was said traditionally that intension determines extension.

Putnam's confusion comes out even more clearly in his final example. He argues as follows:

Suppose you are like me and can't tell an elm from a beech. We still say that the extension of 'elm' in my idiolect is the same as the extension of 'elm' in anyone else's, viz., the set of all elm trees, and that the set of all beech trees is the extension of 'beech' in both our idiolects. Thus 'elm' in my idiolect has a different extension from 'beech' in your idiolect (as it should). Is it really credible that this difference is brought about by some difference in our concepts? My concept of an elm tree is exactly the same as my concept of a beech tree..19

Here again, the conventional extensions of the words 'elm' and 'beech' are those things to which the terms may correctly be applied. And to apply the terms correctly in anything more than contexts like 'An elm is a tree'—i.e., in anything more than trivial contexts—it is necessary to be able to distinguish between elms and beeches.20 This ability, which is not shared necessarily by all native speakers of English, is reflected in the difference in the respective conventional intensions of these words, which reflects, in turn, that someone in the language community can distinguish successfully between elms and beeches. To reiterate, if we distinguish between subjective and conventional intension, we can say that both terms have the same subjective
intension (and subjective extension) for Putnam; and in this (subjective) sense we can say that intension determines extension. Of course, it is wrongheaded to suppose that one's subjective intension determines the (conventional) extension of a word. This is brought out quite clearly by Putnam's examples. But to my knowledge, no traditionalist would wish to be interpreted in this way.

What Putnam finds intriguing, apparently, is that even when one such as I, who cannot distinguish an elm from a beech, makes statements about elms, I am talking about elms and not merely about what I believe to be elms. Therefore, he reasons, the word 'elm' denotes elms in my idiolect. The reason for this, as Putnam well knows, is that the words 'elm' and 'beech' are elements of a common language which I share with many speakers some of whom are able to distinguish elms from beeches. Moreover, as an adult speaker of English I know this, and it is the fact that I know this that enables me to employ such terms. One might want to say that this is proof that I can refer to elms and beeches even though I cannot distinguish between them. But this is highly misleading. I can 'refer' to elms in the sense that I can use the word 'elm' correctly in a rather restricted range of significant sentences. But this is radically different from the sense in which one refers to an elm when one says 'That elm over there (pointing) looks diseased.' In this latter case there is an intentional act of reference involved; to succeed at it the speaker must be able to distinguish the particular elm tree that he intends from others. Of course, it may be the case that the intended tree is not an elm—the speaker is mistaken—but if it is an elm and if the speaker is also able to say (correctly) 'But that beech over there (pointing) is healthy,' then that speaker must be able to
distinguish elms from beeches. We can account for this by saying that
the speaker has different subjective intensions for 'elm' and 'beech'.
Obviously, a speaker who can refer to elms in this latter sense has a
higher degree of competence vis-à-vis 'elm' and 'beech' than a speaker
whose concepts of elm and beech do not differ. Nevertheless, these
differences in the competence of individual speakers have little to do
with whether intension determines extension (in the conventional sense
of 'intension').

Putnam's own positive account of this problem agrees in certain
respects with the above explanation. He postulates the following
"sociolinguistic hypothesis" which he calls the "hypothesis of the
universality of the division of linguistic labor":

Every linguistic community exemplifies the sort of division
of linguistic labor just described; that is, possesses at
least some terms whose associated 'criteria' are known only
to a subset of the speakers who acquire the terms, and whose
use by other speakers depends upon a structured cooperation
between them and the speakers in the relevant subsets.

As an example of this hypothesis Putnam points to our use of the word
'gold':

[e]veryone to whom gold is important for any reason has to
acquire the word 'gold', but he does not have to acquire
the method of recognizing whether something is or is not
gold. He can rely on a special subclass of speakers. The
features that are generally thought to be present in con-
nection with a general name--necessary and sufficient
conditions for membership in the extension, ways of
recognizing whether something is in the extension,
'criteria', etc.--are all present in the linguistic com-
munity considered as a collective body.

We can agree, in large part, with Putnam's hypothesis; yet there are
points with which we must take issue.
The first of these concerns his guarded use of the word 'criteria', hence his use of scare quotes in the quoted passages. This is evidently intended to disassociate his notion of criteria from the traditional notion of conventional intensions. But the distance between the two is illusory: Putnam's invoking of criteria—like possessing the structure H2O for 'water'—differs only verbally from the traditional attempts to characterize meaning in terms of what was contemporarily regarded as the essence of the things in the term's extension. In effect, Putnam's hypothesis recommends replacing Locke's nominal essence with Locke's real essence in cases where the real essence is known. This feature of Putnam's theory makes it difficult to distinguish his view from the more traditional approach.

Secondly, Putnam's claim that our use of, say, 'gold' "depends on a structured cooperation" between lay speakers and metallurgists (who know the real essence of gold) is certainly questionable. This claim appears tantamount to the assumption that the scientific meaning (or use) of a word determines its use in the wider community. But certainly this is doubtful. As Locke pointed out long ago, languages are older than science and because of this the latest scientific discoveries have a less than absolute effect on the established speech habits of native speakers. A more accurate picture, perhaps, would distinguish various uses of the word 'gold', ranging from its application by the unscientific to gold-plated jewelry, to its application by scientists to the white colored substance with atomic number 79. This is not to deny that our use of 'gold' depends, in part, on the latest scientific techniques for recognizing gold, but only to suggest that the dependence is far less
"structured" than Putnam supposes. Putnam concludes his attack on intensionalism with the observation that "it is only the sociolinguistic state of the collective linguistic body to which the speaker belongs that fixes the extension." My claim, and the point of the above discussion, is that if we had put this statement to Locke or Mill or Carnap or Frege, they would have agreed entirely.

Nevertheless, we can credit Putnam with making explicit what was more or less implicit in the traditional view of meaning. This is important and commendable. Unfortunately, Putnam's thesis that natural kind terms are rigid designators—a thesis influenced strongly by Kripke—appears to be at odds with this commendable view in certain important respects. Let us turn to rigid designation then.

**Putnam's Account of Rigid Designation**

The key concept involved in the doctrine of rigid designation is the notion of 'possible world'. Putnam takes it as "primitive" and we shall do likewise (for the sake of argument). Following Carnap a possible world may be represented in terms of a set of sentences that completely describe the world. The ontology of such a world would consist of those objects that must be taken as values of the variables in order to make every sentence of the description true. The best way to deal with Putnam's doctrine is to set out his argument step by step.

He begins by reconsidering his Twin Earth example and asks us to assume that he is giving a meaning explanation of 'water' by "pointing to a glass and saying 'this is water.'" He supposes further that $W_1$ and $W_2$ are two possible worlds corresponding to Earth and Twin Earth
and that the sentence 'This is water' is uttered in each world. In $W_1$ the glass is full of $H_2O$ and in $W_2$ it is full of $XYZ$. Putnam concludes that there are two theories one might have concerning the meaning of 'water' as it is used in these worlds:

(1) One might hold that 'water' was world-relative but constant in meaning (i.e., the word has a constant relative meaning). On this theory, 'water' means the same in $W_1$ and $W_2$; it's just that water is $H_2O$ in $W_1$ and water is $XYZ$ in $W_2$.

(2) One might hold that water is $H_2O$ in all worlds (the stuff called 'water' in $W_2$ isn't water), but 'water' doesn't have the same meaning in $W_1$ and $W_2$. Putnam contends that (2) is "clearly the correct theory." He argues as follows:

When I say 'this (liquid) is water', the 'this' is, so to speak, a de re 'this'—i.e. the force of my explanation is that 'water' is whatever bears a certain equivalence relation (the relation we called 'sameL' above) to the piece of liquid referred to as 'this' in the actual world.

The former of these two theories corresponds to the intensionalist theory and the latter to the rigid designation theory. On the former, the meaning of 'water' is a matter of its intension, which is the same in $W_1$ and $W_2$. (Hence, Putnam's remark that the meaning is "constant" on the first theory.) However, according to Putnam, on the former theory the extensions of 'water' would differ in $W_1$ and $W_2$. As we saw in the preceding section, this is mistaken, for it fails to acknowledge that speakers on $W_1$ and $W_2$ would apply 'water' to both $H_2O$ and $XYZ$ if the conventional intension of 'water' was the same for $W_1$ and $W_2$. In contrast, on the latter theory, the meaning of 'water' is determined extensionally. Putnam explains theory (2) in terms of what we might call an extensional definition. According to this view 'water' designates
whatever is the same liquid as the stuff that is called 'water' in W₁ (the actual world). In other words, according to theory (2), 'water' is a rigid designator; it designates the same stuff in every possible world.

Putnam explains the latter theory as follows:

(2') (For every world W) (For every x in W) (x is water \equiv x bears same L sub W to the entity referred to as 'this' in the actual world, W₁).³³

The key difference between this definition of 'water' and that of the intensionalist theory (1) is that on this theory the sameness L relation is indexed to the actual world; that is, something would be water in some possible world, W₂, if and only if it bears a certain sameness relation to what is called 'water' in the actual world. In contrast, according to Putnam's interpretation of the intensionalist theory, something is water in a world W if and only if it is the same as the stuff called 'water' in that world W. It is easy to see that as Putnam sets out the example 'water' may be applied to different stuff--H₂O or XYZ--in different possible worlds (again assuming the isolation of the respective language communities and employing Putnam's rather implausible notion of extension). So according to the intensionalist view, as Putnam interprets it, 'water' may designate different stuff in different possible worlds. According to theory (2) however, the rigid designation theory, 'water' designates the same stuff in every possible world; that is, it designates H₂O in every possible world.

Putnam explains his 'sameness L' relation as follows:

My 'ostensive definition' of water has the following empirical presupposition: that the body of liquid I am pointing to bears a certain sameness relation (say, x is the same liquid as y, or x is the same L as y) to most of the stuff I and other
speakers in my linguistic community have on other occasions called 'water'.

He qualifies this explanation by arguing that if the empirical presupposition is false—e.g., if the designated liquid turns out to be gin—then "I do not intend my ostensive definition to be accepted." Thus, the ostensive definition gives the necessary and sufficient conditions for something being water only if the stuff in the glass is, in fact, water.

Putnam claims that his sameness relation is a "theoretical relation" and admits that in practice it may take an indeterminate amount of investigation to determine that something is the same as the stuff in the glass. For example, consider again the 1750 version of the Twin Earth example. On Putnam's theory it would make no difference if speakers on Earth and Twin Earth called the stuff in the lakes and rivers of Twin Earth 'water'. At that time they did not know that the stuff called 'water' on Twin Earth does not bear the sameness relation to the stuff called 'water' on Earth. And it would take years of investigation to uncover this fact. But this makes no difference for the extension of 'water' which, unbeknownst to the Earthian speaker of English in 1750, is just that stuff that has the chemical formula H2O. Therefore, even if Earthian speakers consistently called the stuff in the lakes of Twin Earth 'water' that stuff would not be in the extension of 'water'. To my mind, it is highly counterintuitive that a theory of meaning could be thus divorced from how a particular word—or, more generally, a language—is used. To my mind, such a theory must be mistaken.

Part of the counterintuitiveness of Putnam's claim stems from the very improbability of his science fiction examples. We might reasonably ask what a possible world like W2 (Twin Earth) has to do with the
meaning of 'water' as it is used on Earth. It is by now a commonplace to say that meaning is a matter of conventions; by 'conventions' here we mean, of course, conventions that are operative in the language, conventions that are followed (in some sense of this word) by speakers of the language. What, we might ask, does the essence of water——i.e., the essence according to our present-day theories——have to do with the meaning of 'water' as it was used in 1750? Shouldn't its meaning then, as it is now, be a matter of the conventions that were operative then? And since our present-day knowledge of water could in no way have affected the conventions operative in 1750 what does it have to do with the meaning of 'water' as this word was used in 1750?37

There are two ways in which we might interpret the rigid designation thesis——in particular, the ostensive definition on which it hinges. However, the first interpretation leads to absurdity and the second to triviality (unless supplemented with absurdities of its own). Let's take these up in turn. On the first view the extension of, say, 'water', for now and all time, is picked out by the description 'the stuff that has atomic formula H₂O'. Anything that has this formula, or contains elements that have this formula, is water and nothing else should be called 'water' or should have been called 'water'. The absurdity is that the extension of 'water' has nothing to do with how the word is/was used. Moreover, in order to say that a particular application of 'water' is mistaken we must appeal to some criterion or other. In this case, our criterion is 'having atomic formula H₂O'. It might reasonably be argued that this criterion contributes to our present understanding of 'water'. But this is not Putnam's claim. For example, consider his
claim that we should employ present-day scientific knowledge to judge whether Archimedes was applying the word 'χρυσός' correctly in Attic Greek.

We will not distinguish between 'gold' and the cognate words in Greek, Latin, etc. And we will focus on 'gold' in the sense of gold in the solid state. With this understood we maintain: 'gold' has not changed its extension (or not changed it significantly) in two thousand years . . . Let X be [a piece of metal which could not have been determined not to be gold in Archimedes' day]. Clearly X does not lie in the extension of 'gold' in standard English; my view is that it did not lie in the extension of χρυσός in Attic Greek, either, although an ancient Greek would have mistaken X for gold (or rather, χρυσός). . . . But who's to say he would have been wrong? The obvious answer is: we are (using the best theory available today).38

Putnam's remarks in the above passage confuse two questions which are best kept distinct. First, we can now say, in retrospect, that Archimedes was wrong in thinking that X was gold. This is an epistemological question. On the other hand, based on the best theories available in his day—which is all any speaker can reasonably be expected to go on—Archimedes' use of the word 'χρυσός' to refer to X was not mistaken: His use of the term conformed to the conventions governing the use of 'χρυσός' that were operative in Attic Greek. This is a linguistic observation. Questions of meaning cannot be divorced completely from questions of knowledge but the proper relation between these is not what Putnam's theory supposes. Our use of language is governed by what we believe to be true, not by what is in fact true or by what will be held true by speakers 2000 years hence. Thus, given the state of science in his day Archimedes was entirely justified—that is to say linguistically but not scientifically correct—in applying 'χρυσός' to X. His employment of 'χρυσός' would have been mistaken only if X did not conform to the criteria of recognition supplied by the theories then
available. To talk about mistaken applications of a word makes sense only if the speakers of the language can provide reasons (in the form of criteria of applicability) for why a particular application of a word is mistaken. Mistakes in usage, we can say, are intra-linguistic not inter-linguistic, as Putnam supposes. (On the other hand, we can, employing the words of our language in conformity with the conventions that govern their use, make the epistemological claim that Archimedes was wrong in thinking that X was χρυσός--his thinking was wrong, perhaps, but not his use of the Greek language.) In sum, on this interpretation, Putnam's thesis boils down to the view that the extension of a word could be constituted by a criterion of recognition that was in no way available to speakers of the language. We might sum up the fault of Putnam's theory, on this interpretation, by saying that it suffers from epistemological chauvinism.

One of the features that Putnam finds most attractive about his (and Kripke's) view of meaning is its so-called "realist perspective." He claims that such notions as 'truth' and 'extension' or 'truth-of' belong to the realist's idiom and do not fit at all well with the operationalist perspective that attends the intensionalist approach to meaning. Putnam's above claim about the extension of χρυσός follows from this realist perspective. The extension of χρυσός is "just what the term is true of" and truth is here decided by "the best theory available today." That is, as Putnam points out, his theory of language takes truth to be an extra-theoretic notion that is not dependent necessarily on the theories presupposed by users of the particular language under study. This form of realism, I maintain, is
quite misleading and has no role to play in theory of meaning. First, we must revise the statement made in the penultimate sentence: Putnam's theory does not take truth as an extra-theoretic notion; it merely tries to do this but fails. The notion of truth that his theory presupposes is based on the theories that contemporary scientists hold true—e.g., 'Water is H₂O', 'Gold has atomic number 79', and so on. This notion of truth is as intra-theoretic (and intra-linguistic) as the notion that he criticizes the operationalist—i.e., the intensionalist, in this context—for holding. He argues that

\[ \text{the trouble is that for a strong antirealist truth makes no sense except as an intra-theoretic notion. The antirealist can use truth intra-theoretically in the sense of a 'redundancy theory'; but he does not have the notions of truth and reference available extra-theoretically. But extension is tied to the notion of truth. The extension of a term is just what the term is true-of. Rather than try to retain the notion of extension via an awkward operationalism, the antirealist should reject the notion of extension as he does the notion of truth (in any extra-theoretic sense).}^{42} \]

There are several problems with this argument. First, it is difficult to make sense out of a genuinely extra-theoretic concept of truth without invoking a Platonist metaphysics. To see this we might make use of Locke's distinction between real and nominal essence. Locke distinguished between the known nominal essence of things, which include those readily observable properties that speakers (of his day especially) use in recognizing the things, and the unknown real essence, which consists of the microstructure of the thing. There are two ways we might interpret this distinction depending on whether we want to emphasize the essential hiddenness of real essences or the fact that real essence is a matter of the microstructure of things. On the former
interpretation, even if we discover the microstructure of gold we can still talk about a more fundamental real essence that is still hidden from us. On this view our knowledge of the microstructure becomes part of the nominal essence of gold; it affects our use of the word 'gold' in certain contexts. This view borders on Platonism, since real essence becomes some kind of ideal towards which our theories aim but never attain. Moreover on this view truth becomes an ideal. Popper's third world and philosophy of science, and Peirce's theory of truth come readily to mind. Here we have a genuine case of extra-theoretic truth, but it has nothing to do with meaning, since meaning is a function of the nominal essence. On the second interpretation, when we discover the real essence, the distinction between real and nominal essence collapses—as it does, on Locke's theory, for the so-called "mixed modes" like 'bachelor', 'adultery', etc. On this view, truth is intra-theoretic—i.e., current theories are regarded as the standards of truth—and it is not divorced from meaning since real and nominal essence are identical. This interpretation squares with Putnam's assumptions and reveals the mistake in his reasoning: He wants to say we know real essences—i.e., our theories are true—and that real essences are constitutive of meaning, but at the same time, he maintains that truth is extra-theoretic (false) and meaning is extensional (also false). In short, this aspect of Putnam's view is hopelessly inconsistent.

Secondly, it is true that the notion of extension is tied to the notion of truth, but the appropriate connection, as all other extensionalists concede, is between 'extension' and 'true-of-in-L'. That is, the notion of truth employed in the theory of language must be an
intra-theoretic (and intra-linguistic) notion. One's application of
the words of one's language is governed by what one, and others of one's
language community, believe to be true, not by what some language commun-
ity 2000 years hence might believe to be true, nor by what is true in
some Platonist sense of the term. This is so obvious that we needn't
belabor the point. We must conclude then that on his pseudo-realist
interpretation (which is actually theoretical or epistemological chau-
vinism in disguise) Putnam's theory is plainly false. Extra-theoretic
realism may make for good metaphysics—although I seriously doubt it—but it, and Putnam's counterfeit, do not make for very good theory of
language.

At the risk of beating a dead horse we should also point out that
Putnam's avowed realism stands in obvious tension with his sociolinguis-
tic hypothesis. On the one hand he wants to say that "it is only the
sociolinguistic state of the collective linguistic body to which a
speaker belongs that fixes the extension"; on the other hand he wants
to say that the extension of a term "is just what the term is true-of"
where truth is determined by "the best theory available today." If
this isn't plainly contradictory then it is plainly false as the follow-
ing dilemma reveals.

First, the sense in which the two theses constitute an incon-
sistency. If we assume, as apparently we must, that Archimedes is not
a member of the same linguistic community as you, Putnam and I, then it
is certainly possible that the sociolinguistic conventions implicit in
his language community would fix the extension of 'χρυσός' differently
than the sociolinguistic conventions or our community fix the extension
of our word 'gold'. Surely this is possible. Archimedes lived in a culture that had a science and world view vastly different from ours. Surely their science was not as sophisticated as ours, but just as surely there were scientists and hence 'experts' (in Putnam's sense) in Archimedes' day. Like our own experts the scientists of Archimedes' day employed certain theories when confronted by questions like 'Is X a piece of gold?' Similarly, just like our own scientists these experts had definitions like 'The essence of gold is such and such'. Although the 'such and such' had nothing to do with the microstructure of gold it was the best available theory in that day. Moreover, all of these theories affected the use of 'χρυσός', at least in those ranges of discourse that were then regarded as scientific. So it is entirely likely that on the sociolinguistic hypothesis the extension of 'χρυσός' was different from the extension of 'gold'. This plainly contradicts Putnam's realist conclusion that the extensions of both terms are the same. That's one horn of the dilemma.

We can avoid this horn by jumping onto the other. We do this by assuming that Archimedes is a member of our language community. As a member of our community Archimedes and the best scientists of his day were simply misusing the language in saying that X is χρυσός. We can view their assertion on a par with that of a child who is just learning the language and doesn't yet know that it is wrong to apply the word 'gold' to objects that don't have atomic number 79. On this view, the extension of 'χρυσός' is whatever the experts of our society say it is, so its extension is the same as that of our word 'gold'. On this interpretation the realist hypothesis—i.e., that 'χρυσός' and 'gold' have the
same extension—and the sociolinguistic hypothesis—i.e., that extension is determined by the sociolinguistic state of the collective linguistic body—are compatible, but the result is absurd. The problem with this view, obviously, is that no one in his proper lights would want to maintain that Attic Greek and contemporary American English are parts of the same language, or that Archimedes and we are members of the same 'collective linguistic body'. To blur such distinctions would be disastrous for the theory of meaning. The point then is this. We can reconcile Putnam's sociolinguistic hypothesis with his realism only if we blur the differences that exist between historically distinct languages and cultures. This move is absurd. But if we don't blur these differences then the two hypotheses yield contradictory conclusions. We can avoid inconsistency by wallowing in absurdity and vice versa. Clearly Putnam cannot hold both hypotheses. My recommendation is that we drop the (pseudo) realism.

Let us now look at a second interpretation of Putnam's rigid designation theory. On this view the extension of 'water' is not picked out by the description 'the stuff that has atomic formula H2O'; rather, it is picked out by the description 'the stuff that bears a certain samenessL relation to "most of the stuff that speakers in [our] linguistic community have on other occasions called 'water'."' Notice the reference to a 'linguistic community' here. Now, it is obvious (and trivial) that the extension of 'water' for a given language community is whatever bears a certain similarity to the stuff that is called 'water' by members of that community. But clearly this truism won't yield—nor should it yield—the conclusion that the extension of 'χρύσος' and the extension
of 'gold' are the same (in Putnam's sense of 'same'). Thus, this interpretation yields the rigid designation thesis only if it is supplemented with certain other assumptions. These assumptions appear to be that the sameness relation be characterized in terms of certain criteria. This is reasonable enough. But ordinarily whenever we define a word on the basis of similarity, we mean 'same in certain important respects' and these important respects are usually construed as respects that are evident to the language user. But to characterize the sameness relation in this way won't yield the rigid designation thesis either, since the important respects that were evident to Archimedes are different from those that are evident to us. Moreover, to characterize the sameness relation in this way, and in any way whatsoever, is a form of intensionalism in disguise, since the "important respects" function as criteria of recognition in the traditional sense. This poses a problem for the rigid designation thesis: If Putnam characterizes the sameness relation in terms of 'water is H2O' he ends up with a theory that is indistinguishable from intensionalism. In this sense this interpretation collapses into the one considered above—i.e., if we employ such criteria of recognition across languages and cultures as Putnam does. Therefore, Putnam must attempt to obtain the proper fix on the extensions of 'χρυσός' and 'gold' without characterizing the sameness relation in any way. This appears hopeless.

To see that it is, let us look more closely at Putnam's claim that when Archimedes asserted that X is gold (χρυσός), "he was saying that it had the same general hidden structure (the same 'essence' so to speak) as any normal piece of local gold."49 In this case Archimedes has not said
what the essence is—he can't since he doesn't know it—so if he can define 'gold' rigidly in this way, we'd be compelled to say 'Hats off to rigid designation!' However, Locke long ago scotched such false hopes when he pointed out the uselessness of such definitions.50

Employing Locke's distinction between an unknown real essence and a known nominal essence, let's suppose that in Archimedes' day the nominal essence of gold was its yellow color, heavy weight, malleability and fusibility.51 These are the characteristics—the 'stereotype', to employ Putnam's terminology52—that Archimedes would use to recognize gold. Let's suppose that X has all these characteristics but is not made of the element with atomic number 79. Of course, Archimedes would say that X is gold (χρυσός) and correctly so judging by the theories implicit in his language. According to Putnam, when Archimedes does say this he is saying that X has the same real essence "as any normal piece of local gold."53 However, even if he was saying this, which is extremely doubtful, it makes not a shred of difference as regards the meaning of 'χρυσός' in his idiolect. For, neither Archimedes nor any of his contemporaries could know whether any of the things that they called 'χρυσός' (or pointed to when they give their ostensive definition) were really gold in Putnam's sense. So Archimedes could not have used Putnam's sameness relation to give an ostensive definition of 'χρυσός'. It is always possible that Archimedes would pick up a sample that looks like gold in all other respects but isn't real gold.

Putnam would claim that in that instance "[Archimedes] would not intend [his] ostensive definition to be accepted."54 The reason is that the definition works only if Archimedes happens to pick up a piece of
metal that has atomic number 79. With a word like 'gold' the chances are relatively good, perhaps, but this is not true in general. For example, it is extremely doubtful that one could have used the sameness relation to define 'fish' by pointing to a fish before it was discovered that whales, dolphins and porpoises are not fish. Perhaps there is some ad hoc disclaimer that can get us around this difficulty too, but a theory that can work only by adopting such ad hoc assumptions is not very useful in any general way. The point is that the sameness relation is useful only after we know the real essence of gold. Before we know the real essence, Putnam's definition could work only by accident. Locke made much the same point when he remarked that "if we know not the real essence of gold, it is impossible we should know what parcel of matter has that essence, and so whether it be true gold or no." Theories of meaning that work only by accident, we might say, are not very useful theories.

It might be argued that I'm being unfair to Putnam's thesis by rejecting it because it works only by accident. For, it might be argued, it is entirely reasonable to assume that the Attic Greeks were quite able to recognize pieces of real gold. Therefore, there's no good reason to believe that such ostensive definitions would fail uniformly. The problem with this reply is that it trivializes Putnam's thesis: If Archimedes' criteria were sufficient for picking out the correct extension of gold, then there is no reason to appeal to the real essence of gold in order to explain the meaning of 'gold'. The intensionalist is quite prepared to admit that the extension of 'gold' has not changed significantly from the extension of 'χρυσός' as used in Archimedes' day.
But in admitting this he grants nothing to the rigid designation theory, for the intensionalist can easily account for the relative stability of the extension of 'gold' on his own principles: The extension of 'gold' has not changed significantly through the years because the superficial properties of gold have served quite well in identifying samples of gold. We can reasonably assume that the discovery that gold has a certain microstructure has made little difference in our use of the word 'gold'; except for very artificial situations such criteria are just not very useful to language users.\(^{56}\) Locke also pointed this out long ago:

\[\text{But supposing that the real essences of substances were discoverable by those that would severely apply themselves to that inquiry, yet we could not reasonably think that the ranking of things under general names was regulated by those internal real constitutions, or anything else but their obvious appearances; since languages, in all countries, have been established long before sciences.}^{57}\]

Locke's point is compelling; the meanings of our words are determined by conventions that were in existence long before the discovery of contemporary atomic theory. So we needn't appeal to the rigid designation thesis to account for the relative stability in the conventions of our language.

Moreover, what is at issue between the intensionalist and the rigid designation theory is a theoretical question. Therefore, even if we grant that such ostensive definitions sometimes work (by accident) this does not constitute evidence in favor of rigid designation. The reason is that the rigid designation thesis does not just boil down to the claim that the extensions of our words are relatively stable; it makes the additional claim that the extensions of our words (such as 'gold') are the same in all possible worlds. This appeal to 'possible worlds' is
quite essential to the thesis, and it is to this point that the evidence must be directed. In an important respect it is only this further claim that would make a difference in cash value between Putnam's theory and the intensionalist theory for, as we have seen, the intensionalist can quite easily account for the relative stability of the extensions of words. Indeed, the fact that in most cases Putnam's thesis would not yield a difference in cash value from the traditional view is the reason that he must resort to such blatantly fictional examples in order to explain the thesis. But there is a problem, in principle, with such examples (insofar as they are used to draw conclusions about meanings). Namely, if the possible worlds involved in such examples differ very radically from our own world, then we not only end up with a different world, we end up with a different language or, at least, with the same language plus some different semantic conventions. Therefore, it is difficult to see what would even count as empirical evidence in favor of Putnam's thesis.

Consider the Twin Earth example again. First, differences that are not known to the speakers of Earth (W₁) and Twin Earth (W₂) are irrelevant for theory of meaning, since these differences would not show up in sentences held true by the native speakers. Therefore, in 1750 there are no relevant differences between speakers of W₁ and W₂ as regards their use of language. In 1950 however there are differences. 'Water is H₂O' is held true in W₁ while 'Water is XYZ' is held true in W₂. The objects in the extensions of 'waterₑ' are mirror images, so to speak, of the objects in the extension of 'waterₑᵗₑ'--if we assume that the two cultures do not overlap in any way, visit each other, etc. In
this case the languages are distinguishable, however, as would become
evident if speakers of one language visit the planet of the other. For
example, speakers of Earthian English might not apply the word 'water' to
the stuff in the lakes and rivers of Twin Earth--i.e., if they took
'Water is H2O' as their criterion of application for 'water'. Putnam
would partially agree with this conclusion but would add that the example
proves his thesis that 'water', as we use the term, is a rigid designator,
since speakers from Earth would not apply the term to the stuff in the
lakes of Twin Earth. But where does the force of the 'would not' in this
sort of claim derive its validity? Also, what could possibly count as
evidence that speakers of English would not apply the word 'water' to
stuff that looked like water in all respects save for its hidden atomic
structure? There's nothing to prevent speakers of English from applying
the word 'water' to stuff that looked like water but lacked its atomic
structure; and if the word 'water' was used in this way, its extension
would consist of wholes of both H2O and XYZ.

As these considerations show, the notion of possible worlds sheds
little light on the problem of meaning, since there are numerous ways to
interpret the fictional, possible worlds examples: How can Putnam
maintain that a word has the same extension in all possible worlds if it
is open to us to view the other possible worlds as worlds involving
different languages or as worlds involving the same languages but with
some different linguistic conventions? Kripke argues that in talking
about a word denoting the same object in all possible worlds he means a
word "in our language as we use it."58 If the above considerations are
correct then we see that Kripke's caveat is of little help. For, the
problem is that our language does not possess the sort of conventions that would enable us to interpret such counterfactual (possible worlds) examples in a nonarbitrary manner. Putnam and Kripke to the contrary, no amount of talk about rigid designation and possible worlds will alter this fact about our language.

I shall leave off the discussion of rigid designation at this juncture and take up the question in a somewhat different form when we come to discuss Kripke's version of the doctrine. However, I have little faith that much sense can be made of the doctrine as a theory of meaning. Speaker's aren't rigid and neither are languages, and therefore, I see no good reason to assume that the tools of our languages—viz., the words—are rigid in meaning. As we have seen, Putnam's arguments do much to bolster our skepticism, and, as we shall see, Kripke's arguments do even more.
NOTES--CHAPTER III

1. Putnam's views on meaning are set out in a number of articles. The most complete exposition of his view is contained in 'The Meaning of "Meaning"' in Keith Gunderson (ed.), Language, Mind, and Knowledge (Minneapolis: University of Minnesota Press, 1975), pp. 131-193. We shall work primarily from this article as it contains revised and updated versions of arguments that Putnam expressed in earlier articles. Two of these other articles include 'Is Semantics Possible?', Metaphilosophy vol. 1 (1970), pp. 187-201, and 'Meaning and Reference', The Journal of Philosophy vol. 70 (1973), pp. 699-711. Both of these papers are reprinted in Stephen P. Schwartz (ed.), Naming, Necessity and Natural Kinds (Ithaca, N.Y.: Cornell University Press, 1977), pp. 102-118 and 119-132 respectively. Page references to these latter articles are to their reprints in Schwartz.


3. Ibid.

4. In most cases the distinction between being X and being called "X" makes little difference. For example, we call certain things 'human' and we assume that they are human; in our everyday use of language we do not assume that the things we call 'human' may not, in fact, be human. The distinction will prove useful, however, in those cases where we learn subsequently that some of the things called 'X' no longer count as being X. For example, before the discovery that whales are mammals, it is reasonable to assume that they were called 'fish' and correctly so called if we judge by the conventions observed by the community of speakers who did not know that whales are, in fact, mammals. Yet we now know that whales are not fish, so for our language community the extension of 'fish' does not include whales. For that earlier community, however, we must hold that the extension of 'fish' did include whales, since they applied the word 'fish' to whales.


6. Ibid., p. 134.

7. In general outline at least, but perhaps not in all specifics, Putnam would probably not disagree with this exposition of 'knowing the meaning.' For example, he compares understanding a word with "knowing how to factor in one's head [which] is just a matter of being in a certain very complex psychological state." ('The Meaning . . .', p. 135) Both kinds of knowing are capacities.

8. The following examples and quotes are taken from Putnam's 'Meaning and Reference.' The same examples are employed in 'The Meaning . . .'}

10. Ibid., p. 121.

11. Ibid.

12. Ibid.

13. Ibid., p. 122.

14. The linguistic competence of a speaker refers to his ability to produce and understand a potential of infinitude of sentences of his language; his performance consists in his actual performance. As Chomsky points out, "[l]inguistic theory is concerned primarily with an ideal speaker-listener, in a completely homogeneous speech-community, who knows his language perfectly and is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, . . . errors (random or characteristic) in applying his knowledge of the language in actual performance . . . . We thus make a fundamental distinction between competence (the speaker-hearer's knowledge of his language) and performance (the actual use of language in concrete situations)." (Noam Chomsky, Aspects of the Theory of Syntax (Cambridge, Massachusetts: The M.I.T. Press, 1965), pp. 3-4.

15. Putnam, 'Meaning and Reference,' p. 121: "If a space ship from Earth ever visits Twin Earth, then the supposition at first will be that 'water' has the same meaning on Earth and Twin Earth."

16. Ibid., p. 124.

17. I assume here that when those like Locke characterized meaning in terms of 'ideas in the mind of the speaker' what they intended were the communicable content of the ideas, not the ideas as subjective entities or mental particulars. When traditionalists argued that intension determined extension they meant 'intension' in the sense of 'conventional intension'. It is true that they viewed the conventional intension on the mentalistic model, assuming that it was shared by speakers who used the word correctly, but this confusion still does not amount to the view that subjective intension determines extension.

18. It was pointed out to me that 'subjective intension' may not be quite right here and what is needed is a short expression for 'intension shared by a language community unaware of scientific findings'. The point is well-taken. However, we can make do with the notion of 'subjective intension' in these examples if we permit the fiction that Oscar represents the typical lay speaker and that all such lay speakers share the same intension. In this case, then, the intension is subjective only in the sense that it is not shared by all speakers of the language community but rather by a proper subset of the community. This qualification does not
affect the thrust of my argument, however, since the linguist who is interested in the extension of 'aluminum' for the linguistic community as a collective body must take into account the use of 'aluminum' by all segments of the community. As a glance at any good dictionary will reveal, this appears to be what lexicographers do; thus, an explanation of the meaning of 'gold' includes the atomic properties of gold as well as those properties with which the lay speaker is more likely to be familiar.


20. This requires qualification to allow for the fact that speakers who are unaware of any relevant difference between beeches and elms can, so to speak, accidentally refer correctly to an elm with the word 'elm'. But this does not affect my argument since it is very doubtful that such accidents would be repeatable uniformly. Such contingencies must be distinguished from cases where the speaker does not intend to refer to some particular object (which happens to be an elm) but just uses the word 'elm' in a significant sentence like 'Elms are common in New England.' How can such a speaker, who can't distinguish elms from beeches, be talking about elms? This problem is addressed below p. 51-52.

21. Again, this requires refinement to rule out cases where Putnam accidentally applies the words 'elm' and 'beech' to elms and beeches. Perhaps we can say that the applications must be repeatable. The utility of the notions of subjective intension and subjective extension lies in the ability to describe the linguistic competence of an individual speaker by employing such notions. Clearly, there is a difference in Putnam's competence as regards 'elm' and that of the botanist. The proposed refinement would appear justified on these grounds at least.

22. See Michael Dummett, Frege: The Philosophy of Language (London: Gerald Duckworth and Company Limited, 1973), pp. 138-140: "... there is no sharp line between [a case in which one acts as a recording device and uses a name like 'Mr. Cunningham' to relay a message: 'Mr. Cunningham telephoned ...'] and a fully fledged mastery of a name: a whole series of transitional cases stretches from one to the other. If a person knows of Milan only that it is a city somewhere on the continent of Europe, we should hardly ascribe to him a complete grasp of the name 'Milan' ... In a great many such cases we are exploiting the fact, known to us, that the word we use is part of the common language. We use the name of a town, knowing only that it is a smallish town somewhere in southern Spain, secure in the knowledge that it could be more precisely identified by ourselves or our hearers, when necessary, by means of maps, reference works, road signs or questions addressed to people living in the area ... What we are relying on is the fact that the name is part of established usage." The same argument applies mutatis mutandis to Putnam's elm and beech example.

24. Ibid., p. 146.

25. Ibid., p. 145.

26. See above, Chapter II, note 16.


28. Carnap, op. cit., p. 9: "A class of sentences in \( S_1 \) which contains for every atomic sentence either this sentence or its negation, but not both, and no other sentences, is called a state-description in \( S_1 \), because it obviously gives a complete description of a possible state of the universe of individuals with respect to all properties and relations expressed by the predicates of the system. Thus the state-descriptions represent Leibniz' possible worlds or Wittgenstein's possible state of affairs."


30. Ibid.

31. Ibid.

32. Ibid.

33. Ibid., p. 149.

34. Ibid., p. 142.

35. Ibid.

36. Ibid.

37. Donald Davidson has expressed similar doubts about Kripke's thesis but only in passing. See 'First General Discussion Session,' Synthese vol. 27 (1974), pp. 506-507.


39. Ibid., p. 155.

40. By 'operationalism' in this context Putnam is apparently referring to the use of observable characteristics of an object as criteria for recognition of the object.


42. Ibid., p. 154.


46. Actually, as regards the word 'χρυσός' this is not very likely since properties like malleability, fusibility, heavy weight, etc., serve pretty well in picking out gold. Therefore it may well be the case, for reasons that have nothing to do with rigid designation, that the extension of 'χρυσός' was not significantly different from that of our word 'gold'. This point is developed below, pp. 67-68. However, this does not affect my point since the conventions governing the use of 'χρυσός' and 'gold' are different. We might not apply 'gold' to things that are malleable, fusible, etc., but which do not have atomic number 79, but such a problem would not even have arisen in Archimedes' day.

47. To employ Kuhn's lingo, our scientific paradigm vis-à-vis gold has replaced Archimedes'. Kuhn would claim that this change in paradigm brings with it a change in the meaning (both intensional and, less significantly in this case perhaps, extensional) of 'gold': the word is used somewhat differently now that we believe in the atomic theory. See Thomas Kuhn, The Structure of Scientific Revolutions (2nd ed. enl.; Chicago: The University of Chicago Press, 1970), chs. V, IX, X, XI. Also, as Paul Feyerabend has pointed out, whether theory change brings with it changes in the meaning of words depends on whether we focus on the everyday uses of words or their more narrow scientific uses. Take the expression 'rigid body'. According to Newtonian mechanics all bodies are rigid, but according to Einstein's general relativity there are no rigid bodies. Thus the extension of this term differs rather drastically from one theory to the other. So if we look at how this term is used in statements or explanations of the theories, we'd be led to say that the phrase has changed in meaning. On the other hand, there has been little change in the everyday use of the expression. We still talk about the desk I am presently writing on as a rigid body. According to Putnam's theory this would be a mistaken use of the English language since 'rigid body' is not true of this desk. For Feyerabend's argument see his 'On the "Meaning" of Scientific Terms,' The Journal of Philosophy vol. 62 (1965), pp. 266-274.

48. Putnam, 'The Meaning . . .', p. 142. We might characterize the difference between these two interpretations as follows: On the former, the extension of a word is picked out by a certain description—'that stuff with formula H₂O'. To my mind this interpretation differs only verbally from intensionalism since meaning,
in part at least, is a function of that description. Kripke would argue that the description serves only to pick out the reference but is not to be construed as part of the meaning of the term. Keith Donnellan and David Kaplan both think it possible to fix the extension of a term rigidly in this way--i.e., by means of descriptions. I challenge this claim below, arguing that there is, in principle, no way to distinguish fixing the referent of a term, in this sense, from defining the term itself. This comes out very clear when we look at the use to which Kripke puts such extensional definitions. See below p. 139. On the latter interpretation, the extension is not fixed by appeal to any particular description. The problem here is that it's not fixed at all in this way. See the text.


50. Locke, III, vi, 50. John L. Mackie, in Problems from Locke (Oxford: Clarendon Press, 1976), Ch. III, §7, has argued that Locke anticipated but rejected Kripke's rigid designation thesis. This is not quite true. Locke argued that "when a man says 'gold is malleable' he means and would insinuate something more than this, that what I call gold is malleable (though truly it amounts to no more), but would have this understood, viz. that gold, i.e., what has the real essence of gold, is malleable . . ." See III, x, 17-19. Actually what Locke means by this does not coincide exactly with the rigid designation thesis. Locke thought all the kinds that we distinguish in nature are nominal kinds, therefore, strictly speaking, we cannot say that all the things that we call 'gold' are really gold, since we may one day discover fundamental ways in which some objects that we call 'gold' differ from other objects that we call 'gold'. On the other hand Locke recognized that we all behave and speak as if our picture of the world were true. This is true enough. But it is silly to insist that those things that we call 'man' are not really man; such statements make sense only within some kind of Platonist framework. A more reasonable view perhaps is this: Our theories of the world may prove to be false or may one day be abandoned in favor of more fundamental theories. But this observation is true of all conceptual schemes and amounts to saying that we can never be absolutely certain that our theories are true. However, because the same can be said for the world view that was operative in Locke's day and our present world view, such observations are of limited value as regards theory of meaning. For, clearly, our use of language assumes that our theories of the world--in the broad sense of 'theories'--are true, that what we call 'man' are really men, and so forth. Therefore these statements like 'Well our theories could be wrong . . .' drop out of the picture when we come to talk about meaning. We must start with what language users hold true and forget that what they hold true may actually be false. To my mind, similar remarks would apply to questions like whether a word denotes the same extension in all possible worlds; our language is about this world and there's no predicting how we would react, linguistically, if we discover that some of our cherished beliefs--e.g., Water is H2O--turn out to be false.
Locke's argument against the view that our words stand for real essences is that such a thesis would be useless so long as we are ignorant of the real essences. I shall develop this point in the text. Locke suggests other reasons for rejecting this view and it may be worthwhile to consider one of them briefly. It is that such extensional definitions—'that stuff with such and such a real essence'—are useless "when the body itself is away." (Ibid.) This is similar to the point I was trying to make in note 48 above: We can, by accident, define 'gold' by saying 'whatever has the same essence as this (pointing).' Such extensional definitions are of little use however if we don't know the real essence, for we don't then know if this is real gold or not. On the other hand it might be urged that once we know the real essence (or think we know it) we could introduce the term by means of that essence. The problem with this is that we end up with a view that is indistinguishable from intensionalism. Mackie urges that we take this course, and he tries to distinguish it from intensionalism by arguing that "we can introduce the stuff (gold) by way of the criteria of recognition, and yet annex the word by way of them to that stuff and not to the criteria." (p. 97) There could scarcely be anything more arbitrary, more ad hoc and less empty (in empirical cash value) than this proposal.

51. For Locke's distinction see his Essay, III, iii, 15-17 and vi; for a brief account see above p. 13-14.


53. Ibid., p. 153. Putnam often characterizes his rigid designation thesis in terms of what speakers intend or mean or say. Thus he argues that "[t]he rigidity of the term 'water' follows from the fact that when I give the ostensive definition 'this (liquid) is water' I intend [that for every world W and any x in W, x is water if and only if x bears sameL to the entity referred to as 'this' in the actual world W1]." (p. 149) This is an unfortunate way to put the thesis for it is by no means obvious that speakers intend any such possible worlds gobbledy-gook when they give an ostensive definition of a term.

54. Ibid., p. 142.

55. Locke, III, vi, 50.

56. To put this same point somewhat differently, we might say that both 'the stuff with atomic number 79' and 'the metal that is yellow colored, heavy in weight, fusible, malleable' pick out virtually the same extension. This is a common phenomenon. Witness 'creature with a heart' and 'creature with a kidney'.

57. Locke, III, vi, 25.
58. Saul Kripke, 'Identity and Necessity,' in Schwartz (ed.), p. 78. For a more detailed discussion of this claim on Kripke's part, see below p. 127-128.
CHAPTER IV

THE INTENSIONALIST APPROACH TO PROPER NAMES

Fregean Sense and Reference

The intensionalist approach to proper names is characterized by its commitment to what is called 'the principle of identifying descriptions' (PID). Strawson has provided the following statement of this principle:

... it is no good using a name for a particular unless one knows who or what is referred to by the use of the name. A name is worthless without a backing of descriptions which can be produced on demand to explain its application.¹

Strawson's idea is that a proper name is, in some sense of the phrase, associated with a backing of descriptions that serves to pick out or identify the intended referent of the name. For example, if one says 'Socrates was snub-nosed' then the description that could be produced in answer to the question 'Who is Socrates?' would constitute the backing of descriptions; these would pick out the unique individual that was the intended referent of 'Socrates'. Searle has argued that the PID really boils down to "what Frege was getting at ... when he pointed out that a referring expression must have a sense ... or ... a meaning, a descriptive content, in order for a speaker to succeed in referring when he utters it."²

Although the above statement of the principle requires much refinement and qualification, most philosophers have thought that some such principle must be invoked to account for the referring function of proper names. Nevertheless, recently, the principle has been attacked
on several points by Kripke and Donnellan. Both reject the traditional approach and both have proposed alternative explanations of how the referent of a proper name should be determined. We shall consider their objections to the PID and their own theories in due course, but in this chapter we shall be concerned with the PID itself.

In spite of the familiarity of the PID certain questions about its interpretation have nevertheless been raised. For one, in contrast to Searle's view of the matter, Donnellan has argued that the PID "should not be thought of as expressing the [Fregean] thesis that proper names have a sense . . . [because that thesis] . . . suffers from a vagueness about what is to count as showing that an expression has a sense." For another, while it appears clear that the PID purports to be a theory of reference, Kripke has raised the question of whether the PID is to be construed as a theory of meaning or a theory of reference, in Quine's sense of that distinction. Kripke lumps together Frege's thesis that names have sense with Russell's theory of descriptions and argues that both Frege and Russell were mistaken in thinking that proper names could be "defined synonymously" in terms of definite descriptions. As we shall see, it is doubtful that either Frege or Russell held the view that Kripke attributes to them.

Our purpose in this chapter is mainly expository. In the first section, I shall discuss some of the arguments which support the Fregean thesis that names have sense. In the second section, we shall examine the alternative approach, first suggested by Mill. In the third section, we attempt to clarify the claim that a name requires a backing of descriptions. And in the final section, I argue that Kripke's interpretation of the PID is mistaken.
Donnellan's claim that we don't know what is to count as showing that names have sense is somewhat surprising in light of the solid arguments that have been offered in support of this thesis by Frege and Russell. Frege's argument concerns the role of proper names in identity statements. He argues that if our understanding of a proper name consists entirely of the association of a certain referent with it, then there would be no way to explain the difference in cognitive value between statements of the form 'a=a' and 'a=b' (assuming 'a=b' is true). Statements like 'a=a' always hold \textit{a priori} but those like 'a=b' "often contain very valuable extensions of our knowledge and cannot always be established \textit{a priori}." According to Frege, the fact that there are synthetic identity statements involving proper names shows that names must express a sense as well as designate a reference: "[although] the truth value of 'a=b' is the same as that of 'a=a' . . . the sense of 'b' may differ from that of 'a', and thereby the thought expressed in 'a=b' differs from that of 'a=a'." Thus the difference in cognitive value between 'a=a' and 'a=b' is due to the difference in sense between 'a' and 'b'. Thus, Frege's argument is a \textit{reductio} of the view that names lack sense: if names lack sense, there would not be synthetic identity statements.

As an example that lends support to Frege's argument we might take the sentence 'The Morning Star is The Evening Star.' The truth of this statement was discovered (by some Babylonian perhaps) through empirical observation--i.e., its truth was discovered \textit{a posteriori} and, we might also add, its truth could only have been discovered \textit{a posteriori}. As Quine points out, "the two phrases 'Morning Star' and
'Evening Star' cannot be regarded as having the same meaning; otherwise that Babylonian could have dispensed with his observation and contented himself with reflecting on the meanings of his words.\textsuperscript{11}

Russell also presented an argument that is closely parallel to Frege's.\textsuperscript{12} According to Russell, if a sentence like 'Scott is Sir Walter' is not a tautology, it is because the names involved are not being used as names but as descriptions. In that case the sentence means "the person called 'Scott' is the person called 'Sir Walter'."\textsuperscript{13} Russell's reasoning is similar to Frege's: If "to know the meaning of a name is to know who it is applied to" then true identity statements involving names should be tautologous.\textsuperscript{14} The fact that they are not (always) tautologous shows that the meaning of the expressions does not consist merely in their having a referent. Despite these similarities in their reasoning Russell's view differs significantly from Frege's. According to Russell the proper names of our ordinary language are not genuine names at all but "truncated" or "abbreviated" descriptions.\textsuperscript{15} A genuine name--a logically proper name--according to Russell, is a symbol "directly designating an individual which is its meaning .. ."\textsuperscript{16} For example, "the name 'Romulus' is not really a name but a sort of truncated description. It stands for a person who did such-and-such things, who killed Remus, and founded Rome, and so on. It is short for that description .. ."\textsuperscript{17} Russell also offered a second argument for the view that ordinary names are disguised descriptions. Because the meaning of a genuine name is its referent, one cannot sensibly ask 'Does Socrates exist?' nor sensibly assert 'Romulus does not exist' if 'Socrates' and 'Romulus' are genuine names.\textsuperscript{18} Because we can sensibly raise such questions these names are really disguised descriptions.
As I suggested, there are important differences between Russell's and Frege's views of names. Frege did not acknowledge a class of so-called genuine names—logically proper names—whose meanings were their referents; he simply regarded our ordinary proper names as genuine names, insisting, however, that all names have sense—i.e., descriptive content. There are also similarities between their views: Russell's claim that ordinary proper names are really truncated descriptions may be viewed as an acknowledgement of the truth of the Fregean thesis that names have sense; for it is an admission that in some sense such names have a cognitive content that goes beyond their being merely associated with a referent. However, despite this similarity we shall do well to distinguish sharply between their respective theories. This is something that Kripke fails to do, and the results are misleading.

One crucial difference between Russell's and Frege's views concerns the notion of existence. According to Russell's theory, definite descriptions—i.e., expressions of the form 'the so and so'—involve a uniqueness claim and an existence claim. On Russell's analysis the sentence, 'The present king of France is bald' means 'There exists a unique thing x such that x is now king of France and x is bald.'\(^{19}\) Because no object satisfies these conditions the sentence itself is false. Also, because of Russell's 'truncated description' theory, a similar analysis would apply to 'Socrates is snub-nosed'. On Russell's theory 'Socrates' is short for 'the so and so' where 'so and so' represents a conjunction of those properties believed-true of Socrates. If no individual exists that satisfies these properties—or a weighted selection of them—then the sentence 'Socrates was snub-nosed' is false.
On Russell's view then, a sentence like 'Socrates is snub nosed' means 'There exists an x such that x is the so and so and x is snub-nosed.' Frege's theory, on the other hand, is not committed to the view that the use of a name in a sentence involves the affirmation that some individual exists which satisfies a sufficient number of the descriptions associated with the name. His view involves the weaker claim that most uses of proper names presuppose that the name has a referent. If the name lacks reference then we conclude that sentences involving the name are neither true nor false. The same applies for definite descriptions, which Frege regarded on a par with proper names; the use of a definite description presupposes but does not assert that the expression has a referent. The difference between their respective views is not merely terminological nor is it merely one of emphasis, for, as we shall see, Russell's view amounts to the claim that the meaning of a proper name is some set of definite descriptions whereas Frege's view amounts to the claim that a name is associated with a backing of descriptions that determines its referent but does not give the meaning of the name (in any significant sense of 'meaning'). We shall return to this matter below.

Mill's Theory of Proper Names

Despite the differences in their respective views it is safe to say that both Russell and Frege were aware of the fact that our understanding of proper names does not consist entirely of knowledge of their referents; and, Donnellan to the contrary, the existence of synthetic identity statements and significant negative existential statements provides a relatively clear illustration of "what is to count as showing that an
expression has a sense.” Donnellan presents some arguments that provide an alternative explanation of negative existential statements, but we shall postpone discussion of these for the present. For now suffice it to say that the existence of synthetic identity statements and significant existential statements provides at least prima facie grounds for the truth of the claim that proper names have sense.

Let us turn now to the opposite claim and see what sort of arguments can be given to support the view that proper names lack sense. Mill held this view. He argued that "[p]roper names are not connotative [but merely] denote the individual [they name because] . . . they do not indicate or imply any attributes as belonging to those individuals." He compared the use of a proper name with the chalk mark made by the robber in Arabian Nights: "The chalk mark does not declare anything about the house; it does not mean This is such a person's house, or This is a house which contains booty. The object of making the mark is merely distinction." Mill's view then is that names lack intensional meaning and, like the chalk mark, "are attached to the objects themselves . . . "

Mill's best argument in support of this view is perhaps the following:

[A] town may have been named Dartmouth because it is situated at the mouth of the Dart. But it is no part of the signification . . . of the word Dartmouth, to be situated at the mouth of the Dart. If sand should choke up the mouth of the river, or an earthquake change its course, and remove it to a distance from the town, the name of the town would not necessarily be changed. That fact, therefore, can form no part of the signification of the word; for otherwise, when the fact confessedly ceased to be true, no one would any longer think of applying the name. Proper names are attached to the objects themselves and are not dependent on any attribute of the object.
Much of what Mill says in this passage is true; however, it is not particularly relevant to the thesis that names have sense (or connotation). The problem is that Mill here takes the sense of a proper name to have the same logical (or semantical) role as that of the connotations of general terms; but their respective roles are significantly different.

In the first place, it is obvious, as Mill notes, that proper names are not "dependent" on attributes for their application in the same way that predicative terms are. The term 'human' is applicable to certain objects because they satisfy a certain criterion of application, and it would be mistaken to predicate it of objects that fail to satisfy this criterion. But no parallel exists in the case of proper names; we do not apply names to things because they satisfy certain criteria. We give proper names to things so as to be able to call attention to them without having to describe them, and the reasons why we give them a particular name (rather than some other name) are irrelevant. Thus, the sense of a proper name, unlike the connotation of a predicate expression, does not function as a criterion of applicability. Were we interested in such a criterion for proper names, the appropriate question would appear to be: Is the object the bearer of the name? An object comes to be the bearer of a name by being given that name directly--by a christening or a dubbing--or by acquiring the name in some less direct way. But surely it is irrelevant to ask for the attributes in virtue of which it was given the name. This would be like asking why horses were initially given the name 'horse' instead of the name 'dog'. The point is that the sense of a proper name does not purport to explain why an object was given its particular name, nor does it purport to provide criteria of
applicability for the name. Its role is that of identification, not applicability; it purports to answer questions like 'Tom whom are you referring by "X"?' where 'X' is a proper name.

This difference in logical role between the senses of proper names and predicate expressions corresponds to the difference in the logical role between names and predicates themselves. Names and other denoting expressions are employed primarily to pick out or call attention to individuals, while predicates are used to convey information about individuals. Mill argues that names do not convey information about their bearers, and he contrasts them, on this score, with definite descriptions which, he argues, do convey information about their referents and are therefore connotative. However, there appears to be a subtle confusion at work here. The arguments considered above provide prima facie grounds for the view that names have descriptive content. Does this amount to the claim that the use of names conveys information about their referents? The answer, I believe, is 'No'—or, more precisely, names do not convey information when used referentially. But this certainly does not prove that they lack sense for it can reasonably be argued that referential uses of definite descriptions do not convey information about their referents. When I say 'The father of Socrates was not snub-nosed' I use the description 'The father of Socrates' merely to pick out a certain individual. I could also have picked out that individual by using his name: 'Sophronicus was not snub-nosed.' In neither case does the referring expression convey information about that individual in any significant sense of 'convey'. Indeed, I couldn't have
used the description to pick out that individual unless my audience already knew which individual was Socrates' father.

This comes out even more clearly if we look at a description like 'The man with the blue blazer' employed in the presence of the intended referent. Surely it is not the role of the description to convey that some man has a blue blazer; that information is presupposed in the context of its utterance. So it is not clear that we should consider the referring use of a description an instance of conveying information about its referent. Perhaps there is a secondary or degenerate sense in which such uses of denoting expressions convey information. However, two points are in order as regards such a sense of 'convey'. First, in light of Frege's and Russell's arguments, there is every reason to believe that names like 'Homer', 'Socrates', 'Santa Claus' and a wealth of others also convey information in some secondary sense. Secondly, in view of the fact that the primary function of a referring expression is to pick out an individual, we might be well advised to reserve the notion of 'conveying information' for the act of predication for which it does possess a clear meaning.

Mill also argues that predicative uses of proper names do not convey information while predicative uses of descriptions do convey information: "When we predicate of anything its proper name, when we say pointing to a man, this is Brown or Smith . . . we do not merely in doing so, convey to the reader any information about them except that those are their names." But this alleged difference between names and descriptions is quite arbitrary and unfounded. For, in the first place, when we convey the fact that someone is named so and so we just as much "give the hearer . . . entirely new information" as when we predicate some other
attribute of that person. \textsuperscript{29} Secondly, we often use names in quite significant predications as when we say 'He's a real little Napolean' or 'He's no Homer'. In such cases we quite clearly convey information—a wealth of information—which we certainly could not do if such names lacked sense entirely.

Thus, Mill's argument for the thesis that names lack sense is inconclusive. We can agree with Mill that when used referentially names serve merely to pick out the intended referent, but since the same can be said for referential uses of descriptions, this, in itself, does not count against the view that names have sense. Frege's thesis purports to explain how the connection between a use of a name and its referent is effected. Mill's doctrine does not really address this question. However, it is worth noting that insofar as Mill did address this question, the answer he comes up with is very close to the Fregean answer. He argues:

When we impose a proper name, we perform an operation in some degree analogous to what the robber [in the Arabian Nights] intended in chalking the house. We put a mark, not indeed upon the object itself, but, so to speak, upon the idea of the object. A proper name is but an unmeaning mark which we connect in our minds with the idea of the object ... \textsuperscript{30}

This view is only a short remove from Frege's view because, presumably, our idea or concept of an object involves precisely that sort of information that Frege regarded as part of the sense of the object's name. Indeed, on this score, Mill's view is well nigh indistinguishable from Carnap's development of the Fregean view whereby the intension of a proper name is an individual concept. \textsuperscript{31}
The Backing of Descriptions

What then is the sense of a proper name and what role does it play? As Frege explains it, the sense of an expression is that "wherein its mode of presentation [of the referent] is contained." Frege also remarks that "to the sign there corresponds a definite sense and to that in turn a definite reference." These remarks suggest that the sense of an expression provides a means for introducing or picking out or identifying a referent. Moreover, the examples that Frege employs suggest that a name picks out its referent by means of certain descriptive information that is associated with the name. For example, there are a number of different expressions we can use to pick out the planet Venus: 'The Morning Star', 'The Evening Star', 'Venus', etc. Each of these picks out the same object but in slightly different ways because each expression is connected with different—perhaps, by now only slightly different—information about the planet. Moreover, the information with which a name is associated may differ from speaker to speaker and from context to context. According to Frege, with natural languages "one must be content if the same word has the same sense in the same context." Also, the fact that speakers attach different senses to the same name does not prevent speakers "from grasping the same sense." Proof that the senses of names—and other expressions of our language—are communicable resides in the fact that "mankind has a common store of thoughts which is transmitted from one generation to another." Undoubtedly, the relation between this common store of thoughts—we might call them a shared corpus of beliefs—and the senses of our expressions is a symbiotic one: these shared beliefs serve to give a certain
communicability to the expressions of our language which serves in turn to enable us to express and preserve the corpus of beliefs. 39

One question that quite naturally arises in connection with the notion of sense is: What is meant by the claim that the sense determines the referent? The answer is that the sense of a proper name functions as a means for identifying the intended referent. But this needs refinement. It was Frege's view that in order to refer to an object by means of a proper name (or other denoting expression), one must first be able to identify the object (in the sense of being able to distinguish it from others). Thus Frege spoke of a "criterion of identity" (Kennzeichen für die Gleichheit). 40 He employed this notion primarily in talking about references to numbers but with certain modifications the same notion can be used as a general condition for successful reference:

If we are to use the symbol 'a' to signify an object we must have a criterion for deciding in all cases whether B is the same as A, even if it is not always in our power to apply this criterion . . . When we have thus acquired a means of arriving at a determinate number and of recognizing it again as the same, we can assign it a number word as a proper name. 41

Of course, unlike for mathematics, in order to refer to objects in ordinary discourse it needn't be required that one be able to identify an object "in all cases" in order to refer to it in one particular situation; nor need we be able to recognize an object at some future time in order to refer to it successfully in the present. Therefore, what Frege had in mind as identity conditions for reference to mathematical objects is far too rigorous for the purposes of ordinary discourse. Nevertheless, as a general condition for successful reference his point is well taken: In order to refer to an object one must be able to distinguish the object from others in one way or another.
Both Searle and Strawson have adopted this principle as a necessary condition for successful reference. For example, Searle argues that "[i]f a speaker refers to an object, then he identifies or is able on demand to identify that object for the hearer apart from all other objects." This leads Searle to formulate what he calls "the axiom of identification":

A necessary condition for the successful performance of a definite reference in the utterance of an expression is that either the utterance of that expression must communicate to the hearer a description true of, or a fact about, one and only one object, or if the utterance does not communicate such a fact the speaker must be able to substitute an expression, the utterance of which does.

As Searle's remarks suggest, descriptions are expressions that carry their identity criteria on their sleeves, so to speak. However, his formulation of the axiom requires some qualification; for, strictly speaking, it is not necessary that the expression—a definite description, say—communicate "a description true of" the intended referent. As Donnellan has pointed out, one can refer to Jones as 'Smith's murderer' even though it is not true that Jones murdered Smith. However, this does not constitute a serious obstacle for Searle's axiom. We need only modify it to require that the description be believed true of the intended referent. This would suffice to cover most cases where the description employed—either as the expression itself or as the backing for a proper name—is strictly false of the intended referent. For example, it would cover the case where a color-blind speaker refers to a man with red socks on as 'the man with the green socks'. The speaker's reference would succeed if the audience knows of the speaker's color-blindness and adjusts for it. Also, in most such cases the context suffices to guarantee that
the act of reference is successful. For example, the actual description employed may over-describe the referent, so to speak; e.g., in the case supposed, it may be that there is only one individual present with socks on, in which case the reference to their color would be superfluous. Searle is aware of such contingencies and qualifies his thesis accordingly.45

Unlike referring uses of descriptions, names do not wear the identity conditions of their referents on their sleeves, and therefore, they require what Strawson calls a "backing of descriptions." The descriptive backing is what the speaker and his audience would rely on to identify or pick out the intended referent; it may consist either of presumptions about the referent or perceptions of the referent. Strawson explains the situation as follows:

... in any communication situation a hearer (an audience) is antecedently equipped with a certain amount of knowledge, with certain presumptions, with a certain range of possible current perception. There are within the scope of his knowledge or present perception objects which he is able in some way or another to distinguish for himself. The identificatory task of one of the terms, in predications of the kinds we are now concerned with, is to bring it about that the hearer knows which object it is, of all objects within the hearer's scope of knowledge or perception, that the other term [i.e., the predicate term] is being applied to. This identificatory task is characteristically the task of the definite singular term. That term achieves its identificatory purpose by drawing upon what in the widest sense might be called the conditions of its utterance, including what the hearer is presumed to know or to presume already or to be in a position there and then to perceive for himself... The possibility of identification in the relevant sense exists only for an audience antecedently equipped with knowledge or presumptions, or placed in a position of possible perception, which can be drawn on in this way.46

We see then that the PID requires that in most cases the objects referred to by name are identified and distinguished from other objects antecedently. We might say then that the sense of a proper name--its backing
of descriptions—is that information, known to those who have acquired the name, which identifies the bearer of the name. Depending on the name its sense may be fairly well circumscribed and determinate—e.g., names like 'Homer', 'Romulus', 'Santa Claus'—or largely indeterminate—e.g., place names, personal names, etc. It is difficult to generalize about the senses of names but we might try the following: In general, it is doubtful that the sense of a proper name can be exhaustively characterized in terms of descriptions. The reason is that names—especially names for historical and legendary figures—are connected with several descriptions any one of which may suffice to identify uniquely their referents. This does not detract from the usefulness of names—so long as the various descriptions pick out the same object—and may even add to their usefulness. For example, names enable us to acquire new information about their referents without having to describe the referent each and every time we want to talk about it.47

Kripke's Interpretation of the Principle of Identifying Descriptions

I should like to assume at this point that the PID and its connection with Frege's thesis that names have sense is reasonably clear. We shall have ample opportunity to tie up whatever loose ends remain as we discuss recent objections to the principle. For now however let us take up the question, raised by Kripke, as to whether the PID purports to be a theory of meaning or a theory of reference. To begin with, Kripke draws no distinction between Frege's and Russell's view of names but lumps them together as the "Frege-Russell" theory.48 He argues that both Frege and Russell thought that "a proper name, properly used, simply was a definite description abbreviated or disguised."49 Also, he claims that
"Frege specifically said that such a description gives the sense of the name."50 As regards this last claim, Kripke is simply mistaken.51 He cites no passage in which Frege is supposed to have said this, and, to my knowledge, there is no such passage. What Kripke may have in mind are some of Frege's examples, in which Frege does characterize the sense of an expression in terms of definite descriptions. But it must be remembered that these examples are drastically simplified models of the actual sense of a name. In general it is doubtful whether the sense of a name can be completely characterized by a description. This may be possible for names like 'Homer' and 'Romulus'—as Russell's theory suggests—but is not generally the case. As Dummett points out:

What is important about Frege's theory is that a proper name, if it is to be considered as having a determinate sense, must have associated with it a specific criterion for recognizing a given object as the referent of the name. . . . Sometimes the criterion may be capable of being conveyed by means of a definite description, in other cases not. It is therefore of importance to note how much of Kripke's criticism depends upon his excessively narrow interpretation of Frege, and how much is unaffected by taking Frege's theory as a 'sense theory' but not necessarily a 'description theory'. . . .52

As Dummett's remarks suggest, it is misleading to equate Frege's sense theory of names with Russell's description theory. By viewing names as disguised descriptions—i.e., as equivalent to a conjunction of properties believed true of the referent—Russell's theory implies that names have a fixed meaning. While this may be true of some names, it is not true in general. In contrast, Frege's theory is more nearly analogous to Wittgenstein's account of names:

But when I make a statement about Moses—am I always ready to substitute some one of these descriptions for 'Moses'? . . . Has the name 'Moses' got a fixed and unequivocal use for me in all possible cases?—Is it not the case that I have, so to speak, a whole series of props [e.g., definite descriptions] in
readiness, and am ready to lean on one if another should be taken from under me and vice versa?53

In short, the thesis that a proper name has a sense does not commit Frege to the view that the name is replaceable by some particular definite description. On the other hand Russell's claim that names are disguised descriptions does commit him to this view. One way in which we might characterize the difference between their views is to say that Russell's theory purports to give a theory of meaning for names while Frege's purports to give a theory of reference. 54

Kripke argues that as a theory of meaning the Frege-Russell approach would involve the claim that names are "defined synonymously" in terms of descriptions.55 This is definitely mistaken as an interpretation of Frege's thesis and it is doubtful that Russell would have claimed that names and descriptions are synonymous in Kripke's sense. Kripke takes synonymy as substitutivity salve veritate in all contexts, including, if we judge by his arguments, those contexts that Frege called "ungerade."56 For example, Kripke argues that names cannot be substituted for descriptions in sentences like (1):57

(1) The U.S. President in 1970 might not have been the U.S. President in 1970.

For, whereas there is a clear sense in which (1) is true, the sentence which results from (1) by substituting names for descriptions is never true according to Kripke:

(2) Nixon might not have been Nixon.

Kripke regards this phenomenon—the failure of substitutivity in model contexts—as an "intuitive test" of the correctness of his thesis that names (but not descriptions) are rigid designators; i.e., names denote
the same individual in all possible worlds but descriptions may denote differently in different possible worlds.\textsuperscript{58}

We shall take up Kripke's rigid designation thesis below. However, for now, suffice it to note that Frege would not have held that names and descriptions are synonymous in Kripke's sense. Consider the following sentences:

(3) Hegel believed that Aristotle was the teacher of Alexander.

(4) Hegel believed that the Stagirite was the teacher of Alexander.

On Frege's view (4) does not follow from (3) even if we suppose that (5), below, is true:

(5) Aristotle was the Stagirite.

This is so because Hegel might not have known that (5) is true. Frege accounted for the failure of substitutivity in such contexts by arguing that in oblique contexts--of which modal sentences are one example--names denote their ordinary senses not their ordinary referents.\textsuperscript{59} Therefore, (4) does not follow from (3) because 'Aristotle' and 'the Stagirite' denote different objects in that context; each denotes its customary sense. This is tantamount to saying that 'Aristotle' and 'the Stagirite' have different senses, and therefore, they would not be synonymous, but merely coextensive, on Frege's theory. Even this argument falsifies Frege's view somewhat for it is not clear that he viewed the notion of sense as providing a basis for a synonymy relation; indeed, in the case of proper names it is highly doubtful that the notion of synonymy has any bearing in the first place. Therefore, Kripke is quite wrong in attributing to Frege the view that names are synonymous with or defined in terms of descriptions.
Although this is not our main worry here, it is also doubtful that Russell viewed names as synonymous with descriptions in Kripke's sense of 'synonymous'. Russell distinguished between two ways of using ordinary proper names: (i) as genuine names in which case they merely denote an individual, and (ii) as descriptions in which case they are analyzable in terms of definite descriptions. For example, he argued that if the 'C' in 'Scott is C' were replaced by a name—any name for Scott—"then if the name is being used as a name and not as a description, the [resultant] proposition will still be a tautology." (As we shall see, this comes very close to Kripke's analysis of identity statements, whereby he argues that identity statements involving proper names are necessarily true though sometimes a posteriori.) To my knowledge Russell never addressed the question of whether proper names function as names or descriptions in modal contexts, but clearly his distinction, if developed in an appropriate way, would provide a ground for denying that names function as disguised descriptions in such contexts. This is speculation however, and Russell's arguments provide little indication of how he would actually have developed his description theory of names. Nevertheless, at best, the interpretation that names and descriptions are synonymous applies to Russell's and not to Frege's theory of names.

It also would not apply to Searle's theory of names. Searle denies explicitly that names can be said to possess definitions and this should be proof enough that he doesn't intend his theory as a theory of meaning for names. Searle sums up his theory as follows:

Suppose we ask users of the name 'Aristotle' to state what they regard as certain essential and established facts about
Searle also argues that "it is a contingent fact that Aristotle ever went into pedagogy (though I am suggesting that it is a necessary fact that Aristotle has the logical sum, inclusive disjunction, of properties commonly attributed to him."64 In response to this Kripke argues that if Searle is using 'necessary' in the same way that Kripke uses it, then Searle's theory is plainly false; for, "[i]t just is not, in any intuitive sense of necessity, a necessary truth that Aristotle had the properties commonly attributed to him . . . It would seem that it's a contingent fact that Aristotle ever did any of the things commonly attributed to him . . ."65

There are two problems with Kripke's argument. The first concerns the de dicto/de re ambiguity.66 Interpreted de dicto--as the assertion that 'Necessarily Aristotle was $P_1$ or $P_2$ or . . . '--Searle's thesis is quite reasonable. We might symbolize it as follows:

(6) $\Box (P_1(a) \lor P_2(a))$, where 'a' denotes Aristotle and $P_1$, $P_2$ are properties commonly attributed to Aristotle. To keep the example simple we employ a disjunction of two properties.

What this means is that the disjunction of those sentences involving attributions to Aristotle is necessarily true. However, it does not follow from this that each of the disjuncts is necessarily true. That is, it is not the case that (7):

(7) $\Box (P_1(a) \lor P_2(a)) \rightarrow (\Box P_1(a) \lor \Box P_2(a))$

Therefore, it is quite proper for Searle to claim on the one hand that the disjunction is necessarily true and to claim at the same time that
Aristotle might not have gone into pedagogy and might not have been the Stagirite and so on. The claim that (6) is quite compatible with (8):

(8) \( (\sim \Box P_1(a) \cdot \sim \Box P_2(a)) \)

Thus his theory, interpreted de dicto, squares with Kripke's observation that "it's a contingent fact that Aristotle ever did any of the things commonly attributed to him . . ."\(^67\)

The case is vastly different if Searle's thesis is construed de re -- i.e., in the sense 'Aristotle was necessarily Plato's pupil or necessarily the Stagirite . . .'. Assume, for example, that '\( \Box P_1 \)' and '\( \Box P_2 \)' correspond to two such necessary attributes of Aristotle. Then on the de re interpretation Searle's thesis may be symbolized as follows:

(9) \( \Box P_1(a) \vee \Box P_2(a) \)

And of course (9) is incompatible with the claim, interpreted de re, that Aristotle might not have been \( P_1 \) and might not have been \( P_2 \). That is, (9) is incompatible with (10) which is logically equivalent to (11):

(10) \( \Diamond \sim P_1(a) \cdot \Diamond \sim P_2(a) \)

(11) \( \sim \Box P_1(a) \cdot \sim \Box P_2(a) \)

Thus, if interpreted de re, Searle's thesis would be inconsistent for it would conflict with Searle's claim that "it is a contingent fact that Aristotle ever went into pedagogy."\(^68\)

The question then is whether Searle intends the de dicto or de re interpretation, and this leads to the second problem involved in Kripke's argument: He has attributed to Searle a theory of meaning when it is quite obvious that Searle's theory of names is intended as a theory of reference. What Searle appears to mean by his remark that "it is a
necessary fact that Aristotle has some disjunction of the properties commonly attributed to him . . ." is that unless some statements about Aristotle were held true, we could not be said to be referring to a unique individual when we employ the name 'Aristotle'. Searle's claim, as he makes reasonably clear in his exposition, is just a development of Frege's insight that it is a necessary condition of successful reference that one must have picked out or identified the intended referent in some way or other. Frege's Sinn and Searle's disjunction of properties serve to identify or fix the referent of 'Aristotle'. Thus Searle argues that

[i]t is a necessary condition for an object to be Aristotle that it satisfy some of [the descriptions associated with the name 'Aristotle']. This is another way of saying that the disjunction of these descriptions is analytically tied to the name 'Aristotle'--which is a quasi affirmative answer to the question 'Do proper names have sense?' . . . [in the sense that] names are logically connected with characteristics of the object to which they refer . . .

Two points should be noted concerning Searle's claim. First, the claim that descriptions are "analytically tied" or "logically connected" to names should not be viewed as the claim that names are defined in terms of descriptions which, as we have noted, Searle explicitly denies. Therefore, it does not follow from Searle's view, properly interpreted, that, e.g., 'Aristotle is the teacher of Alexander' is an analytic statement. The logical connection between names and descriptions is that the associated descriptions determine the referent of the name. This is quite within the realm of the theory of reference. Secondly, one quite reasonable way to interpret the 'logical connection' between names and descriptions is in terms of sentences-held-true-within-a-language-community. Unless some sentences about Aristotle are held true--i.e.,
unless some distinguishing features of Aristotle were believed true of Aristotle—it is difficult to see how one could be said to be referring to a unique individual when employing the name 'Aristotle'. Both Kripke and Donnellan have denied this, but this is another question (involving their causal or historical theory of reference) and will be taken up in a subsequent chapter. For now, suffice it to say that if the theory of identifying descriptions, interpreted as a theory of reference, is true, then it certainly appears that Searle's claim must be interpreted de dicto: The disjunction of some sentences held true about Aristotle is a necessary condition for successful reference to Aristotle. Viewed thus, Searle's thesis appears prima facie plausible.

In view of these considerations, it is doubtful that Kripke's rigid designation theory can be viewed as a refutation of the intensionalist view, even if his theory is true which, I hope to show, it is not. But Kripke claims that if the intensionalist theory is not interpreted as a theory of meaning, it loses much of its "attractiveness" and is unable to handle successfully the problems which it is intended to solve. This claim is also doubtful. The problems which Kripke has in mind—problems of synthetic identity statements, vacuous referring expressions, negative existential statements—are problems of reference, not problems of meaning, and can be handled quite satisfactorily within the intensionalist theory of reference. It will not be possible to go into all of these problems here, but it should suffice to indicate the general direction which a theory like Searle's would take if we consider the problem of negative existentials.

Kripke expresses some misgivings about his interpretation of the intensionalist view, but he argues, tentatively, that "if 'Moses' means
the same as 'the man who did such and such' then to say that Moses did not exist is to say that the man who did such and such did not exist, that is, that no one person did such and such." This is mistaken. It does not follow from the PID that to say 'Moses does not exist' implies that no one did such and such. Consider Searle's solution to this problem. Searle argues that to say 'Moses did not exist' is tantamount to saying that a sufficient number of the descriptions commonly associated with 'Moses' are false; i.e., such statements as 'Moses led the Israelites out of Egypt' are false. However, it cannot be assumed that in asserting the falsity of 'Moses led the Israelites out of Egypt' we are thereby claiming that 'Someone led the Israelites out of Egypt' is also false. To claim that a sufficient number of the descriptive backings of 'Moses' are false, is to claim only that we haven't picked out a unique referent from the name 'Moses'. One might conceivably disagree with Searle's analysis of this problem, but clearly, it is an analysis of a supposed problem of reference, not an analysis of the meaning of 'Moses'.

We see then that it is misleading to view the PID and Frege's sense thesis of names as theories of meaning; both are theories that purport to explain how the use of a proper name refers to a unique individual. Kripke acknowledges that both theories may be viewed as theories of reference, and both he and Donnellan have provided several arguments which claim to show that the PID is false as a theory of reference. We shall take up these arguments in Chapter VII, but before we do that, let us take up Kripke's rigid designation theory, a theory, we
should remember, that is designed to refute the claim that the meaning of names are just definite descriptions.
NOTES—CHAPTER IV


4. Saul Kripke, 'Naming and Necessity', in D. Davidson and G. Harmon (eds.), pp. 258-260. Kripke distinguishes between two sorts of definition: those used to "give the meaning" and those used to "fix the reference" of a name (p. 274); he then argues that Frege's and Russell's view may be interpreted as providing the former kind of definition--i.e., the cluster of descriptions associated with a name is "synonymous" with the name (pp. 276-277). As we shall see, this appears to be a questionable interpretation of Frege's theory and perhaps even of Russell's theory. For Quine's distinction see Chapter I, note 7.

5. Ibid., p. 278. Kripke calls this the stronger version of the Frege-Russell theory: "the name is simply defined synonymously, as the cluster of descriptions." The weaker version maintains only that the cluster of descriptions determines the referent of the name.


7. Ibid., p. 56.

8. Ibid., p. 78.


10. See Frege, op. cit., p. 57.


14. Ibid., p. 244.

15. See 'Logical Atomism', p. 243 and Mathematical Philosophy, p. 179.


20. Frege, op. cit., p. 61: "... when we say 'the Moon', we do not intend to speak of our idea of the Moon, nor are we satisfied with the sense alone, but we presuppose a reference." Similarly, Frege argues that "anyone who seriously took the sentence 'Odysseus was set ashore at Ithaca while sound asleep' to be true or false would ascribe to the name 'Odysseus' a reference, not merely a sense; for it is of the reference of the name that the predicate is affirmed or denied." (p. 62).

21. Ibid.; see the preceding note.

22. By 'on a par' here I mean that Frege did not distinguish between ordinary proper names and definite descriptions. On Frege's view a proper name is any expression that "has as its reference a definite object." (p. 57) Given Frege's ontology this would mean that what we commonly call proper names, definite descriptions, and even declarative sentences are proper names in Frege's sense.


25. Ibid.
26. Ibid.
27. Ibid.
28. Ibid.
29. Ibid.
30. Ibid.
31. Rudolf Carnap, *Meaning and Necessity* (2nd ed., enlarged; Chicago: The University of Chicago Press, 1956), pp. 39-42. Actually Carnap distinguishes his theory of intension and extension from Frege's theory of sense and reference. However, as Carnap points out, his pair of concepts and Frege's correspond in extensional contexts but not in oblique contexts; unlike Frege, Carnap does not maintain that names have their customary sense as their reference in oblique contexts and argues that Frege's theory leads to an infinite number of entities (both senses and referents). (pp. 129-133) The key theoretical differences between the two are that according to Carnap, Frege's pair of concepts are intended as an explication of the name relation whereas his pair of concepts are intended as an explication of the traditional connotation/denotation distinction. (pp. 126-127). Nevertheless, Carnap's individual concept is enough like Frege's sense to permit the statement in the text.

33. Ibid., p. 58.

34. Michael Dummett explains Frege's sense as follows: "The sense of a word . . . constitutes the contribution which it makes to determining the truth-conditions of sentences in which it occurs precisely by associating a certain reference with it . . . [The sense of a proper name tells us] how [the] association [between a name and its referent] is established. For the purposes of logic, this is unnecessary: for the purposes of a theory of meaning, it is essential" (*op. cit.*, p. 93). For the most part Dummett's interpretation appears unobjectionable. It squares with the Fregean principle that the sense of a name determines its referent. However, there is one matter over which we might rightfully quarrel with Dummett. It is his claim that Frege's notion of sense is essential for a "theory of meaning." I shall argue below that Frege's thesis is best viewed as a theory of reference, not a theory of meaning. This appears justified on the grounds that it is doubtful that Frege's notion of sense can suffice as a basis for synonymy, definition and the other concepts that Quine includes in the theory of meaning. Also, Dummett argues that it was Frege's view that although the notion of reference is a part of the theory of meaning, "the reference of a term is no more part of what is ordinarily understood as its meaning than the truth-value of a sentence is." (p. 84).
This is a very puzzling claim that, I believe, must be false both as an interpretation of Frege and as an analysis of meaning. First, Frege originally thought that meaning was reference and his argument in 'On Sense and Reference' is an acknowledgement that meaning cannot consist entirely of reference. But it doesn't follow from this that he concluded that reference was "no part of the meaning" of a term. It is more likely that Frege viewed both sense and reference as contributing to the meaning of a term. Certainly, a grasp of both goes into our understanding of language. Secondly, if we exclude reference from the meaning of a term, we end up viewing meanings as merely lexical; i.e., the meaning of a word is just some other expression of the language. On this picture our language would resemble an uninterpreted calculus, a most unreasonable analogue. In short, so long as we want to keep to the idea that language is 'about the world' it will not do to exclude reference from the meaning of our expressions.

35. For example Frege argues that "[for] 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 a name: the teacher of Alexander the Great was born in Stagira. So long as the reference remains the same, such variations of sense may be tolerated . . ." (op. cit., p. 58). Incidentally, this argument provides additional grounds for rejecting Dummett's claim that "reference is no part of the meaning of a term." (See preceding note.) For if cognitive meaning were just the sense of an expression, it is difficult to see how communication can take place when speakers attach a different sense to the same word.

36. Frege, p. 58.

37. Ibid., p. 60.

38. Ibid., p. 59.

39. It might be added here that it is difficult to get away from the view that the sense of an expression also includes an indication of its grammatical category; surely this is involved in our understanding of a proper name. Indeed, it appears likely that when we encounter a name we have never heard before, we understand merely that as a name it is a symbol that purports to refer to a unique individual; it could be argued that this understanding derives from a knowledge of the grammatical and semantical role of the expression. Therefore, in addition to taking the sense as that which determines the referent, we must also view it as contributing to our knowledge of the grammatical category of the expression. See Dummett, p. 55 for more on this point.

41. Ibid.

42. Searle, p. 79.

43. Ibid., p. 80.

44. Keith Donnellan, 'Reference and Definite Descriptions', in Stephen P. Schwartz (ed.), *Naming Necessity and Natural Kinds* (Ithaca: Cornell University Press, 1977), p. 47f. Searle also acknowledges this phenomenon: "Sometimes the descriptor may not even be true of the object referred to and yet the reference is successful. Whitehead offers a good example: speaker, 'That criminal is your friend,' hearer, 'He is my friend and you are insulting.'" (Searle, p. 89; The reference is to Whitehead's *The Concept of Nature* (Cambridge, 1920), p. 10).

45. For other qualifications of his thesis see Searle, pp. 88-91.


49. Ibid., p. 255.

50. Ibid.

51. This has been pointed out by Leonard Linsky in *Names and Descriptions* (Chicago: The University of Chicago Press, 1970) pp. 42ff, and Michael Dummett, *op. cit.*, p. 110.

52. Dummett, pp. 110-111.


54. For a characterization of Quine's distinction between theory of meaning and theory of reference see above, Chapter I, note 7.

55. Kripke, p. 278.

56. Generally speaking contexts that are 'ungerade' are those contexts that Quine calls referentially opaque, viz. those contexts for which substitutivity of synonyms and quantification are not admissible. Belief statements, indirect discourse, modal contexts are examples. See Frege, 'On Sense and Reference', pp. 66ff. and W.V.O. Quine *Word and Object* (Cambridge: The M.I.T. Press, 1960), pp. 144-151.

58. See Kripke, 'Identity and Necessity', in Schwartz (ed.), p. 83. Also see Ch. V for a discussion of Kripke's rigid designation thesis.


60. Russell, 'Logical Atomism,' p. 246.

61. Ibid., p. 245.

62. Searle, p. 166.

63. Ibid., p. 169.


65. Kripke, 'Naming and Necessity,' p. 287.

66. As we shall use it throughout this paper the de dicto/de re distinction may be explained as follows: If an entire sentence occurs within the scope of a modal operator—as in 'Necessarily Aristotle was a teacher'—we say the sentence is de dicto; but if the modal operator occurs within the sentence—as in 'Aristotle was necessarily a teacher'—we say the sentence is de re.

67. Kripke, 'Naming and Necessity,' p. 287.


69. Ibid.

70. Ibid., pp. 169-170.


72. Ibid.

Rigid Designation and Formal Semantics

Kripke's thesis that proper names are rigid designators is directed against Russell's theory that names are disguised descriptions. As Kripke explains it, a rigid designator is "a term that designates the same object in all possible worlds" in which it has a referent.¹ Kripke's argument against Russell's abbreviated description theory of names may be put simply as follows: Proper names cannot be abbreviated descriptions because descriptions, but not names, may designate different objects in different possible worlds.² We might illustrate this claim by considering the description, 'the U.S. President in 1970'. As things stand, this description designates Nixon; but things might have gone differently and Nixon might have lost the 1968 election. In the possible world—or counterfactual situation—in which Nixon lost that election, 'the U.S. President in 1970' would designate someone else; e.g., it might designate Humphrey.³ In this chapter, we shall examine Kripke's rigid designation thesis. My key contention is that the notion of rigid designation provides a useful recommendation for how one might interpret proper names for the purpose of formal semantics, but that it fails as a theory of ordinary proper names.

As the reference to 'possible worlds' in Kripke's definition suggests, his notion of rigid designation is closely tied to his formal semantics for quantified modal logic. It will be helpful, therefore, to begin our discussion by looking briefly at his formal semantics.⁴
To begin with, consider the system M which contains formulae A, B, C... that are built out of atomic formulae P, Q, R... by means of the connectives •, ~, and □. In addition to the standard axioms and rules for the propositional calculus, the axiom schemata and rules of M are the following:

A1 □ A → A
A2 □ (A → B) → (□A → □B)
R1 A, A → B/ B
R2 A/□ A

As Kripke points out, if we add the schema, □A → □ □ A, we get Lewis' S4; if we add A → □A + B to M we get the so-called Brouwersche system; if we add □ A + □A, we get Lewis' S5. 

A semantics for modal logic uses the notion of a model structure (m.s.); this is an ordered triple (G, K, R) where K is the set of all possible worlds, R is a reflexive relation on K and G is the real world. As Kripke explains it,

K is the set of all "possible worlds"; G is the "real world". If H₁ and H₂ are two worlds, H₁RH₂ means intuitively that H₂ is "possible relative to H₁"; i.e., that every proposition true in H₂ is possible in H₁. Clearly, then, the relation R should be reflexive; every world H is possible relative to itself, since every proposition true in H is, a fortiori, possible in H. 

As Kripke points out, additional requirements—such as transitivity, symmetry and equivalence—may be imposed on R in order to provide model structures for S4, Brouwersche and S5. These additional requirements correspond to the various "reduction axioms" of modal logic.

In addition to the notion of a model structure we need the notion of a model:

[A] model assigns to each atomic formula (propositional variable) P a truth value T or F in each world H ∈ K. Formally, a model on a model structure (G, K, R) is a binary function φ(P, H) where P
varies over atomic formulae and \( H \) varies over elements of \( K \), whose range is the set \( \{ T, F \} \).^8

The assignment of truth values to the truth functional, non-atomic formulae proceeds in the standard fashion. To illustrate how the assignment of truth values works for the modal operators, we can consider the truth value assignment for propositions of the form \( \Box A \). As Kripke defines it: 

\[
\phi(\Box A, H) = T \text{ iff } \phi(A, H') = T \text{ for every } H' \in K \text{ such that } H H'; \text{ otherwise, } \phi(\Box A, H) = F. \]

Intuitively, this means that a proposition of the form \( \Box A \) is true in \( H \) if and only if \( A \) is true in every world \( H' \) possible relative to \( H \).

For the introduction of quantifiers we must associate a domain of individuals with each world. Intuitively, these are the individuals that exist in that world. A quantificational model structure (q.m.s.) is a model structure \( (G, K, R) \) plus a function \( \Psi \) which assigns to each \( H \in K \) a set \( \Psi(H) \), called the domain of \( H \). We add to the symbols of \( M \) an infinite list of individual variables \( x, y, z \ldots \), and a list of \( n \)-adic predicate letters \( P^n, Q^n \ldots \) for each non-negative integer \( n \). (The superscripts are eliminated when they can be understood from the context.)

To define a quantificational model we suppose that in each world a given \( n \)-adic predicate determines a certain set of ordered \( n \)-tuples, viz. its extension in that world. The extension of \( P(x) \) in \( H \), for example, is those individuals the predicate is true-of in \( H \). Kripke adopts the convention that a statement involving free variables has a truth-value in each world for every assignment to its free variables. Formally this goes as follows:
Let \( U = \bigcup_{H \in K} \psi(H) \). \( U^n \) is the \( n \)th Cartesian product of \( U \) with itself. We define a quantificational model on a q.m.s. \((G,K,R)\) as a binary function \( \phi(P^n,H) \), where the first variable ranges over predicate letters, for arbitrary \( n \), and \( H \) ranges over elements of \( K \). If \( n = 0 \), \( \phi(P^n,H) = T \) or \( F \); if \( n > 1 \), \( \phi(P^n,H) \) is a subset of \( U^n \). For an atomic formula \( P^n(x_1,\ldots,x_n) \), where \( P^n \) is an \( n \)-adic predicate letter and \( n \geq 1 \), given an assignment of elements \( a_1,\ldots,a_n \) of \( U \) to \( x_1,\ldots,x_n \), we define \( \phi(P^n(x_1,\ldots,x_n),H) = T \) if the \( n \)-tuple \((a_1,\ldots,a_n)\) is a member of \( \psi(P^n,H) \); otherwise, \( \phi(P^n(x_1,\ldots,x_n),H) = F \), relative to the given assignment. 10

Finally, for a modal logic with identity we can define \( x = y \) to be true in a world \( H \) when \( x \) and \( y \) are assigned the same value and otherwise false. 11

To see how the notion of rigid designation figures into Kripke's semantics we must consider a formal language with individual constants \( a, b, c, \ldots \), and definite descriptions \( (\lambda x)(\phi x) \), which may be translated 'the \( x \) such that \( \phi x \). Assume—and here we depart from Russell's convention in Principia 12—that descriptions are part of the primitive notation of the language and not introduced contextually as abbreviations. The difference between individual constants and descriptions is spelled out by Kripke as follows:

[In the formal semantics of modal logic, the sense of \( t \) is usually taken to be the (possible partial) function which assigns to each possible world \( H \) the referent of \( t \) in \( H \). For a rigid designator, such a function is constant ... In the formal semantics of intensional logic, suppose we take a definite description to designate, in each world, the object satisfying the description. 13]

We see then that a rigid designator, in the formal semantics for quantified modal logic, is a term that would designate the same individual in all possible worlds in which it has a referent; if a rigid designator \( t \) is assigned the referent \( a \) in the actual world \( G \), then it would designate \( a \) in every world \( H \) in which \( a \) exists. In contrast, definite
descriptions may designate different objects in different possible worlds, since the description may be satisfied by different objects in different worlds.

It may be useful to look at a couple of examples. We shall employ an example suggested by Leonard Linsky. Further, to illustrate Linsky's point that Kripke's thesis about proper names reduces to the claim that names do not induce de dicto/de re ambiguities, we shall consider both the de dicto and de re readings of our modal sentences. Consider then the modal sentence (1):

\( \square \Psi(\exists x)(\Phi x) \)

(We remind the reader that \( \square p \) may be defined in terms of the possibility operator '◊' as follows: \( \square p = df. \sim \Diamond \sim p. \) ) In its de dicto sense (1) can be symbolized as (2) below where the bracketed expression is the scope indicator; if it occurs to the right of the modal operator this indicates that the description occurs within the scope of the modal operator:

\( \square [(\exists x)(\Phi x)] \Psi (\exists x)(\Phi x) \)

In order for (2) to be true, \( \Psi(\exists x)(\Phi x) \) must be true at each possible world. To evaluate \( \Psi(\exists x)(\Phi x) \) at a world \( H \) we must evaluate both '\( \Psi(x) \)' and '\( (\exists x)(\Phi x) \)' at \( H \); if the referent of the description is in the extension of the predicate letter '\( \Psi(x) \)' then '\( \Psi(\exists x)(\Phi x) \)' is true at \( H \). The procedure is slightly different for the de re interpretation of (1) which we symbolize as (3):

\( [(\exists x)(\Phi x)] \square \Psi(\exists x)(\Phi x) \)

In this case, because the description occurs outside the scope of the modal operator, we needn't reevaluate it at each possible world. At each possible world it would have the referent— if it has reference
at all in that world—that it has in the actual world. We do reevaluate
the predicate letter '$\Psi(x)' at each world however. Thus '$\Psi(\exists x)(\phi x)' is
true at a world if the referent of '$\exists x)(\phi x)' in the actual world is a
member of the extension of '$\Psi(x)' in H. In effect then, the definite
description functions here as a rigid designator; it designates the
same object at each possible world when it occurs in de re position.17

Intuitively, sentences like (4),

(4) The U.S. President in 1970 might not have been President in 1970,
are sentences which are interpreted de re. Thus (4) is true if there
is some possible world in which Nixon—the referent in the actual world
of 'the U.S. President in 1970'—was not President in 1970. We might
compare (4) with its de dicto counterpart (5).

(5) Possibly the U.S. President in 1970 was not President in 1970.
(5) is false, since at any possible world, H, the referent of 'the U.S.
President in 1970' will be in the extension of 'U.S. President in 1970';
so there is no possible world in which 'the U.S. President in 1970 was
not U.S. President in 1970' is true. This illustrates Linsky's contention
that modal sentences involving descriptions may induce scope ambiguities—
or may induce de dicto/de re ambiguity—since (4), on Kripke's semantics,
would be true, while (5) turns out false.

To see why modal sentences involving names do not induce such
ambiguities on Kripke's thesis, consider the following example in which
'A' is a rigid designator.

(6) $\square [A] \Psi(A)$
(7) $[A] \square \Psi(A)$

(6) is the de dicto and (7) the de re reading of $\square \Psi(A)$. Since 'A'
is a rigid designator it designates the same object in each possible
world. Suppose it designates $a$. Then (6) would be true if at each
world $H$, $a$ is in the extension of $\psi(x)$ at $H$. Similarly, (7) is true
at $H$ if $a$ is in the extension of $\psi(x)$ at $H$. Thus, if 'A' is a rigid
designator, the truth conditions for (6) and (7) are the same; the truth
conditions of $\psi(A)$ will be the same at each possible world—viz.
whether $a$ is in the extension of $\psi(x)$. Therefore, no scope ambiguities
exist in modal sentences involving rigid designators.

Rigid Designation as a Theory of Ordinary Proper Names

If Linsky's interpretation of Kripke's thesis is correct, then one
way we might test Kripke's thesis is to ask if English modal sentences
involving proper names admit of such scope ambiguities. Dummett has
adduced an example which purports to show that they do induce such
ambiguities. His example employs the name 'St. Anne' which he takes
as a likely candidate for a name that is associated with a single
definite description—viz. 'the mother of Virgin Mary'. Because the
name is tied to this description, Dummett claims that it is quite
correct to say 'St. Anne cannot but have been a parent'. That is, this
would be a true sentence in English. Since it is of the form
$\Box P(St. Anne)$, in order for it to come out true by Kripke's semantics
there can be no possible world in which 'St. Anne' designates an in-
dividual that is not in the extension of 'is a parent' (in that world).
Quite obviously however, Kripke's semantics cannot provide an interpre-
tation under which this sentence comes out true for, on his view, there
is a possible world in which St. Anne is not a parent. Kripke would
argue that 'St. Anne' designates St. Anne at each possible world and that
there is a world in which St. Anne is not in the extension of 'is a
parent'. Thus, for Kripke, it would be true to say 'St Anne might not have been a parent' which directly contradicts Dummett's above claim.

There appears to be no way to settle this disagreement since it boils down to a disagreement over the truth or falsity of certain English sentences—e.g., sentences like 'St. Anne cannot but have been a parent'. According to Dummett's intuitions this sentence is true but according to Kripke's it is false. Moreover, in each case a formal semantics can be provided to back up the conflicting intuitions. If we treat 'St. Anne' as a rigid designator, then 'St. Anne cannot but have been a parent' comes out false; but if we treat it as an abbreviation for 'the mother of Mary', we can easily provide an interpretation under which it comes out true.20

What then are we to make of Dummett's counterexample? Does it refute Kripke's thesis? To my mind the answer is 'No', but the reason it does not refute Kripke's thesis is because Kripke's thesis is just not the sort of thesis that is open to refutation by such examples. Dummett is entirely correct in holding that if we take 'St. Anne' as an abbreviation for 'the mother of Mary', the sentence 'St. Anne is the mother of Mary' (or, 'St. Anne is a parent') is necessarily true; it is analytic.21 Moreover, it is still open to Dummett to hold that 'St. Anne might not have been a parent' is also true. There is no inconsistency here. The former claim is the de dicto claim that a certain English sentence is necessarily true—it follows from the way we have defined 'St. Anne'. The latter is a de re claim that St. Anne has the property of being a parent contingently—she might have remained a virgin. We can make this same de re claim by employing the description,
'the mother of Mary', as follows: 'The mother of Mary might never have become a parent'. And, as we noted above, a semantics can be provided to back up Dummett's intuitions on this score. On the other hand, if, following Kripke, we construe 'St. Anne' as a rigid designator, the sentence 'St. Anne is the mother of Mary' is contingently true. Thus, the dispute boils down to whether the English sentence 'St. Anne is the mother of Mary' is contingent or necessary, and since one may have equally strong intuitions one way or the other—depending on how we construe 'St. Anne'—there is no way to settle this dispute without begging the important questions. Therefore, in effect, neither Kripke's thesis that names are rigid designators nor Russell's thesis that names are abbreviated descriptions can be refuted by appeals to our intuitions about the truth of certain English sentences. This is the basis of my claim that both theses are best viewed really as alternative recommendations for formal procedures of treating names; each proposal may lead to a different semantics and either semantics may be useful—and one more useful than the other—for certain purposes. But it appears idle to ask which recommendation gives us the correct account of ordinary proper names.

Our claim then is that Kripke's thesis that names are rigid designators—or Russell's alternative thesis, for that matter—is an interesting proposal for how we might interpret proper names for the purpose of formal semantics, but it does not provide a theory for how proper names are used in ordinary discourse. We shall now attempt to lend further support to this claim.

One argument that Kripke employs in support of his view is the following:
... proper names are rigid designators for although the man (Nixon) might have been the President, it is not the case that he might not have been Nixon (though he might not have been called 'Nixon').

One way in which we might interpret this argument is to take it as making the claim that the sentence (8),

(8) Nixon might not have been Nixon,

is never true. According to Kripke, (8) is never true because 'Nixon' designates the same individual in every possible world (in which Nixon exists). On the other hand, according to Kripke, the truth of (9),

(9) Nixon might not have been the U.S. President in 1970,

hinges on the fact that in some possible world 'the U.S. President in 1970' designates someone other than Nixon—e.g., it might designate Humphrey if we suppose that Nixon lost the 1968 election—while 'Nixon' still designates Nixon in that possible world—since we are talking about Nixon when we say (9). Kripke argues that comparison of modal sentences like (8) and (9) provide us with an "intuitive test" for the thesis that names are rigid designators:

If we apply this intuitive test to proper names, such as for example 'Richard Nixon', they would seem intuitively to come out to be rigid designators. First, when we talk even about the counterfactual situation in which we suppose Nixon to have done different things, we assume we are still talking about Nixon himself... [Second,] we cannot say 'Nixon might have been a different man from the man he in fact was' unless of course we mean it metaphorically.

As Linsky notes, "Kripke employs this kind of argument repeatedly, and it is his main positive argument for the thesis that ordinary proper names are rigid designators."26

There are several problems with Kripke's "intuitive test". First, his argument seems to be more concerned with the identity of the object
referred to than with how the object is designated. For example, in
the passage leading up to the intuitive test, Kripke talks about the
impossibility of the number nine being a different number than it is.\textsuperscript{27} And his intuitive test itself points to the impossibility of Nixon being
a different person than he is. But it is not immediately obvious that
this sort of consideration has anything to do with designation, rigid
or otherwise. It amounts really to the claim that any object is
necessarily self-identical, and this claim has little to do with how the
object in question is designated. Given the appropriate identity con-
ditions, we can say of any object whatsoever, regardless of whether we
choose to designate it by its name or by a description, that it cannot
have been a different object than it is. Similarly, we cannot truly say,
for example, of the pen with which I am writing, that it may have been
a different pen than it is. In this instance I have referred to the pen
via a definite description, but my remarks appear to satisfy Kripke's
intuitive test for rigid designation. Certainly it does not follow that
'the pen with which I am writing' is a rigid designator, since it may
designate different pens in different possible worlds. Thus, Kripke's
test does not appear to be very helpful in distinguishing rigid from
non-rigid designators.

A second problem with Kripke's argument concerns the inaptness of
the comparisons upon which it is based. For example, he asks us to com-
pare sentences like the following:\textsuperscript{28}

(10) The inventor of bifocals might have been someone other than the
man who \underline{in fact} invented bifocals.

(11) Ben Franklin might not have been Ben Franklin.
The comparison is inapt, because unlike 'being the inventor of bifocals', the property 'being self-identical' is not a property that one can acquire. Therefore it is not surprising that the former property is contingently true of an individual while the latter is necessarily true of the individual.\(^{29}\) Thus, the relevant difference between (10) and (11) appears to lie in our institutions about the nature of the properties attributed to the individual and not in how the individual is designated.\(^{30}\)

If we employ Russell's theory of descriptions we can make this same point in terms of the logical form of (10) and (11). On Russell's theory, (10) and (11) would have different logical forms which can be represented symbolically as follows:

\[(10') (\exists x)(Bx \cdot (y)(By \cdot y=x) \cdot \Diamond \neg Bx) ('Bx' is short for 'inventor of bifocals'.)\]

\[(11') (\exists x)(Fx \cdot (y)(Fy \cdot y=x) \cdot \Diamond (x=x) ('Fx' is short for 'x is Franklin'.)\]

In (11), as (11') shows, we are asserting the possibility of an object being non-self-identical.\(^{31}\) As our previous argument showed, it is always false to assert the lack of self-identity of an object, and this does not depend on whether we name or describe the object. To see this we can compare (11) with (12).

(12) The inventor of bifocals might have been a different person than he in fact was.

We can symbolize (12) as (12'):

\[(12') (\exists x)(Bx \cdot (y)(By \cdot y=x) \cdot \Diamond (x=x) ('Bx' is short for 'inventor of bifocals'.)\]

Both (11') and (12') involve the denial of self-identity and both are false; since in one case we employed a name and in the other a description, nothing hinges on whether the object is named or described. However, in (10), as its symbolization in (10') makes clear, we are not
denying that some object is self-identical; we are asserting that 'being the inventor of bifocals' is a contingent property of the inventor of bifocals. 32 But again, nothing appears to hinge on how we designate the inventor of bifocals; we could have expressed the same proposition by saying (13) which may be symbolized as (13'):

(13) Ben Franklin might not have been the inventor of bifocals.
(13') B(B.F) • ◻¬B(B.F)

In short, whatever initial plausibility we find in Kripke's argument appears to hinge on the inaptness of comparing sentences that could be viewed as having different logical forms; when interpreted in the way suggested, the alleged differences between names and descriptions vanish. The point then is this: To establish the rigid designation thesis it will not suffice to show merely that descriptions can designate different objects in different possible worlds—so much is true but irrelevant; for, as the above analysis shows, we can adequately account for our semantical intuitions about English modal sentences without assuming that proper names are rigid designators. To establish this thesis Kripke must show that proper names cannot designate different objects in different possible worlds.

One way in which Kripke might attempt to do this is to argue that it is never true to say English sentences like (14):

(14) Nixon might not have been Nixon.

As I understand it this is the import of Kripke's claim that "it is not the case that Nixon might not have been Nixon (though he might not have been called 'Nixon')." 33 Thus far I have granted that (14) is false and argued that its falsity hinges on the fact that it is necessarily true
to assert self-identity of any object. Actually, however, as an English sentence, (14) is ambiguous between the following two readings: (14') Nixon might not have been called 'Nixon'.

(14'') Nixon might not have been self-identical.

Kripke admits that we can truly say (14') so it is obvious that he does not take (14) in the sense of (14'), but rather in the sense of (14''). This squares with our above observation that his argument seemed to be drawing on the uncontroversial fact that--in some sense at least--every object is self-identical. However, as was noted above, it was Russell's theory that an ordinary proper name 'N' is sometimes used in the sense of the description, 'the person named "N"'. It follows that on Russell's theory (14) could be used in the sense of (14'). That is, on Russell's theory both (14) and (14') mean the same thing, and either could be rendered as (15) without altering their meaning:

(15) The person called 'Nixon' might not have been called 'Nixon'.

However, if (14) and (15) are synonymous, as Russell's theory implies, then Kripke's intuitive test fails; for (15) is a true sentence, as Kripke himself admits. In short, Kripke to the contrary, there is a sense in which we can truly say that Nixon might not have been Nixon; and, more importantly, in order to provide an interpretation for that sense of (14) in which it is true, we must treat 'Nixon' as a non-rigid designator--i.e., as short for 'the person called "Nixon"'.

Kripke has anticipated this sort of counterexample to his thesis and argues as follows:

Sloppy colloquial speech, which often confuses use and mention, may, of course, express the fact that someone might have been called, or not have been called, 'Aristotle' by saying that he
might have been, or not have been, Aristotle. Occasionally, I have heard such loose usages adduced as counterexamples to the applicability of the present theory to ordinary language. Colloquialisms like these seem to me to create as little problem for my thesis as the success of the 'Impossible Mission Force' creates for the modal law that the impossible does not happen.\

This response to the objection is not very compelling. In the first place, the above counterexample does not appear to hinge on a confusion of use and mention: We are there using the description 'the person named "Nixon"' in place of the name 'Nixon' and vice versa. It strikes me that we might view the above argument as a version of Quine's thesis that proper names are dispensable in favor of predicates like 'Pegasizes', the point being that we can use predicates like 'person who is X' and 'person who is called "X"' to do the same work as names. Predicates such as these are co-extensive with the respective proper names.

Secondly, it is Russell's thesis that a name like 'Nixon' is short for the description 'the person named "Nixon"' and it follows from this that there is a sense in which (14), above, is true. To counter this claim Kripke must provide independent grounds for believing that (14) cannot be understood in Russell's sense. To appeal to "sloppy colloquial speech" does not suffice as a way out for Kripke so long as he intends his thesis as a theory about the use of names in ordinary discourse. Ordinary language is colloquial speech, sloppy or otherwise.

We see then that the dispute between proponents of Kripke's and Russell's theories of names boils down to divergent intuitions about the meaning and truth value of certain English sentences--e.g., (14). Since a semantics can be provided to support either's intuitions, there appears to be no reasonable way to settle such a dispute. In light of this
fact, it is misguided to view either Kripke's rigid designation thesis or Russell's disguised description thesis as theories about how names function in ordinary language; both theses provide us with different formal interpretations of proper names and may therefore be judged in terms of their utility for one purpose or another. But it appears idle to insist that one thesis, rather than the other, provides us with the correct theory of proper names. Our point then is analogous to a point made by Quine in a slightly different connection:

I find the phrase 'logical analysis' misleading in its suggestion that we are exposing a logical structure that lay hidden in the sentence all along. This conception I find both obscure and idle. When we move from verbal sentence to logical formulas we are merely retreating to a notation that has certain technical advantages. . . .

Response to an Objection

It might be argued that the thesis that 'Nixon' is an abbreviation of 'the person called "Nixon"' is a trivialization of Russell's abbreviated description theory. In response to this it must be admitted that we are here concerned with a weak version of Russell's thesis; however, this is in keeping with our purpose which throughout has been to cast doubt on Kripke's claim that names are rigid designators. Our case, up to this point, has been that names cannot be rigid designators since they may be replaced by descriptions like 'the person named "X"' which do not designate rigidly.

Chandler has pointed out an important distinction in connection with Kripke's notion of rigid designation. He notes that it is a contingent fact that we use the name 'Agnew' to designate Agnew. We could use the name 'Ford' to designate Agnew: "That is to say there is a
possible world in which we call Agnew 'Ford' . . . " \textsuperscript{41} Chandler argues that this is irrelevant to Kripke's notion of rigidity and proposes a distinction "between speaking with respect to some possible world, and speaking in such a world . . . " \textsuperscript{42} He argues that "[p]resumably, a 'rigid designator' when used with respect to various possible worlds, always designates the same object." \textsuperscript{43} It might be claimed that my above argument rests on a confusion of this distinction since the possible world which we must imagine in order to provide an interpretation for the true sense of (14), above, is just a possible world in which 'Nixon' is used to designate someone other than Nixon. Kripke himself comes close to making this point when he argues, in a slightly different connection, that when a description is said to designate non-rigidly, it is assumed that the description is an expression "in our language as we use it." \textsuperscript{44}

To my mind, however, this sort of reply is unsatisfactory. First, the reply—and Chandler's distinction—provides Kripke with an ad hoc argument that makes his theory virtually immune to counterexample: A possible world in which Nixon is not called 'Nixon'—in which, say, Humphrey is called 'Nixon'—is ruled out as illegitimate or irrelevant. On the one hand this is unobjectionable; it is unobjectionable if Kripke's rigid designation is viewed as a convenient technical device for interpreting certain (but not all) English modal sentences involving proper names. As such it would fail to provide us with a semantics that give the truth conditions for the sense in which 'Nixon might not have been 'Nixon' is true. To circumvent this apparent failure we might rule out that true sense of this sentence by labelling it "sloppy colloquial
usage." This too would be unobjectionable so long as we do not try to
pass off rigid designation as a theory about ordinary uses of proper
names. Or, we might admit that the above sentence is an exception to the
thesis that proper names are rigid designators and grant that it is an
instance in which a name functions as an abbreviation for a definite
description. On the other hand, if Kripke's rigid designation thesis is
a theory about ordinary uses of proper names, the above ad hoc argument
appears to be totally unjustified.

To see this we might consider the following example. An alternate
way of saying 'Nixon might not have been Nixon', in the sense in which
it is true to say this, would be to say 'There is a possible world in
which Nixon is not called "Nixon"'. Similarly, we can also say, 'There
is a possible world in which the person we call "Nixon" is not called
"Nixon"'. My question is this: What sort of grounds—aside from direct
stipulation—can be adduced to rule out such possible worlds?

Chandler's distinction bears some initial promise on this score.
The answer it provides is that it is a contingent fact that we use
'Nixon' to designate Nixon. But this is no help really for when the
full implications of rigid designation are spelled out, this appears to
be what Kripke must deny: To say that 'Nixon' designates Nixon in all
possible worlds in which it has a referent is to deny that there is a
possible world in which 'Nixon' designates, say, Humphrey. And this is
no different, really, from claiming that 'Nixon' designates Nixon
necessarily. In short, the irrelevance of that possible world in which
'Nixon' designates someone other than Nixon is a rather roundabout way
of saying that Kripke's rigid designation thesis is true by definition; such a thesis, of course, will not pass muster as a theory about ordinary proper names.
1. Saul Kripke, 'Identity and Necessity', in Stephen P. Schwartz (ed.), Naming, Necessity and Natural Kinds (Ithaca, N.Y.: Cornell University Press, 1977), p. 78. In future references to this article I will employ the abbreviation 'IN'. The qualifications 'in those possible worlds where the designator has a referent' or 'in those possible worlds in which its referent exists' are often left unsaid by Kripke. We shall sometimes do likewise in future statements of his thesis.

2. Leonard Linsky sums up Kripke's argument in the claim that proper names behave differently than definite descriptions in modal contexts. See Leonard Linsky, Names and Descriptions (Chicago: The University of Chicago Press, 1977), p. 45. Linsky argues as follows: "The most illuminating interpretation of what Russell means when he says that ordinary proper names are 'truncated' or 'disguised' descriptions is that the logical behavior of such names is just like that of definite descriptions. In Principia Mathematica there are symbols which represent logically proper names, and expressions for definite descriptions are introduced by contextual definition; but there are no symbols representing ordinary proper names. The reason is that Russell thought that ordinary proper names could always be represented in his formal system by descriptions without producing distortions in the logical behavior of those names." Linsky concludes that if Kripke's rigid designation thesis of names is correct, "Kripke will have taught us something very important about the logic of proper names which goes beyond what the extensional logicians could reveal." (p. 50). While it may be useful to characterize the issue between Russell and Kripke as Linsky has done—i.e., in terms of a difference in "logical behavior" between names and descriptions—there are certain aspects of this characterization that are misleading and from which, therefore, I wish to disassociate my argument. Chief among these is the question of to what extent the notions of rigid or non-rigid designation can be said to constitute the "logical behavior" of a referring expression. As I understand it—and as Russell and Frege would likely concur—the logical (or semantical) role of a referring expression is to refer to a unique individual. (For the sake of argument we ignore plural and indefinite referring expressions.) Grasping the meaning of a referring expression is, in great part, a matter of grasping to whom or to what it is used to refer. For Frege, our ability is aided by the sense of the expression, of course, and for Russell the case is much the same for ordinary proper names but not for logically proper names. To my mind, it is doubtful that our grasp of the role of a referring expression involves knowledge of whether it may designate a different object in different possible worlds; moreover, it is doubtful that Kripke would insist that it does involve such knowledge. This being the case, Linsky may have wrongly drawn the issue by characterizing it in terms of different views about the logical role
of proper names. Kripke's rigid designation is largely a technical notion whose prime importance rests with the interpretation of modal sentences given that in these modal sentences, the referring expressions have their customary logical role—viz. that of designating a unique individual.

3. See Kripke, 'Naming and Necessity', in D. Davidson and G. Harman (eds.), Semantics of Natural Language (Dordrecht-Holland: D. Reidel Publishing Company, 1972), pp. 264-265. Future references to this article will employ the abbreviation 'NN'.


5. For a concise discussion of these different modal logics see G. E. Hughes and M. J. Cresswell, An Introduction to Modal Logic (London: Methuen and Co., Ltd., 1968).

6. Kripke, 'Semantical Considerations . . .,' p. 84.

7. Ibid. Also see Hughes and Cresswell, pp. 77ff. for a discussion of the relation R. As they explain R, which they call an 'accessibility relation', it corresponds in M, S4 and S5, to different senses of 'conceivability.' Where R is reflexive we have the weakest sense of conceivability.

8. Kripke, 'Semantical Considerations . . .,' p. 84.

9. Ibid.

10. Ibid., p. 87.

11. Ibid., p. 90.

12. In Principia Mathematica (3 vols., 1st ed.; Cambridge: Cambridge University Press, 1910-1915), Whitehead and Russell introduced definite descriptions contextually, in terms of the following definition: Where '(1x)(Fx)' is short for 'the x that F' s' we have (1x)(Fx) = df(E x)(Fx)(y)(Py+y=x)) That is, for Whitehead and Russell definite descriptions did not constitute part of the primitive symbolism of the language of Principia.


15. Ibid., pp. 56-57. For an explanation of de dicto/de re see Chapter IV, note 66 above.

16. See Linsky, p. 58: "... we keep the value of $(\exists x)(\neg x)$ fixed at its value at $W$ and trace that object to each other possible world to see if it does or does not fall under the extension of $\neg$ at each of these worlds."

17. See Linsky, p. 58. As he puts it, "[t]he semantical role of the scope operator, then, is to determine whether the contained description is to be treated as a rigid or a nonrigid designator."


19. Linsky (p. 55) concurs with Dummett on this point: "This is a good example in my case anyway because I remember the occasion upon which I learned the name. I was looking at Leonardo's famous painting of the Virgin Mary and St. Anne and I asked my wife, 'Who is St. Anne?' The reply was, 'St. Anne is the mother of the Virgin Mary.' To this day that is all I know about St. Anne. For me, the suggestion that St. Anne existed but was not the mother of Mary is simply unintelligible, as it should be in Russell's theory."

20. As Hughes and Cresswell point out (p. 197), Church's formalization of Frege's semantics could provide such a semantics. See Alonzo Church, 'A Formulation of the Logic of Sense and Denotation,' in Paul Henle, H. M. Kallen and S. K. Langer (eds.), Structure, Method and Meaning. New York: The Liberal Arts Press, 1951).

21. Kripke would disagree here. He would argue that such sentences are examples of the contingent a priori. We shall take up this aspect of his theory in the next chapter.

22. Kripke, NN, p. 270.

23. It is not entirely clear in what way the notion of rigid designation affects the falsity of (8). As we have seen, in the formal semantics for modal logic, rigid designation appears to be a matter of stipulation. There are no logical constraints against viewing any expression--e.g., a definite description--as a rigid designator. In its symbolic representation (8) has the form $\diamond \neg(N=N)$ and according to Kripke's semantics it would be true if $\neg(N=N)$ were true in some possible world. But the fact that there is no world in which $\neg(N=N)$ is true appears to be due more to the fact that $\neg(N=N)$ is an instance of the logically false sentence $(x)\neg(x=x)$ than to the fact that 'N' is a rigid designator. We shall have more to say on this point below.

24. Again, there is a problem of how we are to interpret (9)--viz. what is its logical form? One alternative is to view it as an identity
statement: \( \Diamond \sim(N=(lx)(Px)) \) where 'N' is short for 'Nixon' and '(lx)(Px)' is short for 'the x who was President in 1970'. We can then show that (9) is true by showing that there is a possible world in which \( \sim(N=(lx)(Px)) \) is true. Assume \( W_b \) is such a world. In that world 'N' designates Nixon and '(lx)(Px)' designates, say, Humphrey. So '(lx)(Px)' designates non-rigidly since it designates Nixon in the actual world. This, perhaps, is how Kripke would interpret and evaluate (9). A second alternative, one that is more in line with Russell's theory, is to take 'the x that was President in 1970' as a quantified expression, i.e., as \((Ex)(Px\cdot(y)(Py\cdot y=x))\). On this interpretation, (9) would have the following form: \( \Diamond (Ex)(Px\cdot(y)(Py\cdot y=x)\cdot N\neq x) \). Actually, this is not yet the fully analyzed form according to Russell's theory since the name 'N' is also an abbreviated description. When fully analyzed its form would be the following: \( (Ex)(Ey)(Fx\cdot (z)(Fz+z=x)\cdot Ny\cdot (z)(Nz+z=y)\cdot x\neq y) \). There is no problem providing a semantic interpretation of this form of (9): it would be true if there is a world, \( W_b \), such that the individual which uniquely satisfies the predicate 'Px' is different from the individual that uniquely satisfies the predicate 'Ny'. We need make no assumptions about the rigidity (or lack of rigidity) of the predicate symbols. Again, the truth conditions for sentences of this form are those which, in general, apply to all identity statements; the truth conditions do not appear to hinge on whether the designators involved are rigid.

25. Kripke, IN, pp. 81-82.

26. Linsky, p. 61.

27. Kripke, IN, p. 82. Kripke argues as follows: "We can say, for example, that the number of planets might have been a different number from the number it in fact is. For example, there might have been only seven planets. We can say that the inventor of bifocals might have been someone other than the man who in fact invented bifocals. We cannot say, though, that the square root of 81 might have been a different number from the number it in fact is, for that number has to be 9."

28. Kripke, IN, pp. 82-83.

29. We assume here for the sake of argument that it makes sense to talk of necessary and contingent properties at all.

30. See Dummett, pp. 128ff. for more on this point.

31. An alternative symbolization of (11) is as follows: \( (EX)(Fx\cdot(y)(Fy\cdot y=x)\cdot \Diamond \sim Fx) \). This symbolization does not involve the assumption that the falsity of (11) hinges on the fact that it asserts the possibility of an object being non-self-identical. Rather, it interprets (11) as having the same logical form as (10). Kripke would undoubtedly prefer this symbolization (or interpretation)
since it would support his theory if we attach the additional con-
dition that 'x is Franklin', unlike the 'x is the inventor of
bifocals' in (10), is a rigid designator. (For a discussion of this
interpretation of predicates like 'x is Franklin', see note 39
below.) The fact that (11) may plausibly be symbolized (and hence
interpreted) in two different ways, indicates that English sentences
like 'Ben Franklin might not have been Ben Franklin' are ambiguous.
I take up this point on pp. 125 ff. below and show that it leads to
problems for Kripke's rigid designation theory of proper names.

32. In other words, whereas (12) has the form of an identity sentence,
(10) does not. Kripke's comparison, therefore, is inapt. He should
be comparing sentences like (11) and (12), not sentences like (10)
and (11).

33. Kripke, NN, p. 270.

34. See above p. 83.

35. Indeed, I believe it possible to argue for an even stronger point;
viz., the proper interpretation of a sentence like 'Nixon might
not have been Nixon', is as synonymous with 'Nixon might not have
been named "Nixon"'. The point here is that names, when used in
predicate-position, do not normally function as short for 'being
Nixon', but as short for 'being named "Nixon"'. The proof of the
pudding can be seen by considering the following well-known joke:
An airline stewardess, in trying to serve a passenger his coffee,
taps him gently on the shoulder and asks, "Are you sleeping?" The
startled passenger replies, "No, I'm Kowalski." Such jokes would
be possible only if predicates like '___ is Nixon' were normally
interpreted as '___ is named "Nixon"'.


218-224, and Quine, From a Logical Point of View (2nd ed. rev.;
discusses this same possibility but concludes that such descrip-
tions as 'the x such that x Socratizes' are themselves rigid
designators. (See Linsky, p. 54). To my mind, this appears simply
to be false: To the extent that 'being X' is a matter of 'being
called "X"'—an extent which appears to be quite great if we judge
by how these predicates are used in everyday language—the Quinean
predicates must be non-rigid, since an object may be called by
different names in different possible worlds.

38. W.V.O. Quine, 'Methodological Reflections on Current Linguistic
Theory', in Davidson and Harman (eds.), p. 451. I am not alone in
this conclusion. Keith Donnellan has argued that "[i]f we were con-
cerned to know about an actual case whether a name had been intro-
duced, by means of a description, as a rigid designator rather than
as an abbreviation, I am inclined to believe we would be in some
difficulty. Not only is there no conclusive argument of the sort described, I doubt that we could rely in general on linguistic intuitions." ('The Contingent A Priori and Rigid Designation,' Midwest Studies in Philosophy, 2 (1977), p. 15). The 'argument of the sort described' to which Donnellan refers is an appeal to various modal sentences employing names and descriptions--i.e., an appeal to usage.

39. Linsky so argues and claims that Russell's theory may be salvaged in this way. He points to Quine's use of the predicate 'x Socratizes' as such a trivialization of Russell's theory: "The predicate 'x Socratizes' was introduced by Quine for just this purpose in extensional logic and it works as well in modal logic. It is a rigid designator which designates the property of being Socrates. If 'Socrates', as a name, is a rigid designator, so is the predicate 'x Socratizes' which has within its extension in each possible world (in which it has anything within its extension) just that individual who is Socrates." (p. 54.) Linsky's assumption that such descriptions are rigid designators is unwarranted for the following reason: For any given language community the properties 'being Socrates' and 'being called "Socrates"' are coextensive. This is not to say that Socrates was named "Socrates" or originally called "Socrates"--Socrates was given a Greek name back in the 5th century B.C. whose English translation is "Socrates"--but that the person we call "Socrates" and the person who is Socrates are one and the same. This is not to make any grand metaphysical claim about Being but is merely to take notice of how our language functions. In using the name "Socrates" we simply do not entertain the possibility--for it makes little sense to do so--that Socrates is not the person whom we call "Socrates". It follows therefore that if there is a possible world in which the person we call "Socrates" was not called "Socrates", then in that world Socrates would not be Socrates. For example, imagine ourselves back in Greece of the 5th century B.C.; in that world--i.e., at that time--Socrates was not called "Socrates" and it would make little sense to say in that world that Эуфраст is Socrates (assuming that at that time Socrates was called Эукрат). This is not to say that a person's being depends on what he is named but to say, rather, that in describing a person as being so-and-so we cannot get away from using the name by which he is in fact called. Therefore, if 'being called "X"' is a non-rigid designator--as, of course, it is--then so is 'being X'. The only way to avoid this conclusion is to take 'being Socrates' as applicable across possible worlds and to do so we must appeal to certain necessary and sufficient conditions of applicability. There appear to be no such conditions for, given any such criterion Y, we can always say truly, 'Socrates might not have been Y'.

40. With a slight modification of the stronger version of the abbreviated description theory it can also be made to suffice as an alternative to the rigid designation theory. The problem with the stronger version is that while it is true to say 'Necessarily Homer is Homer' it is not true to say 'Necessarily Homer is the author of the Iliad' for
the fact that Homer wrote the *Iliad* is a contingent fact--i.e.,
replacement of a name by the description it abbreviates is not
truth-preserving. The problem here appears to be that we believe
Homer wrote the *Iliad* and this belief forms the basis of our decision
to use 'Homer' as an abbreviation for 'the author of the *Iliad*'. But
this belief, even if true, has nothing to do with the truth con­
ditions for the sentence 'Homer is the author of the *Iliad*': We
believe this to be a true sentence but its truth (or falsity) is
independent of our beliefs; similarly, for the sentence 'Necessarily
Homer is the author of the *Iliad*'. The problem then is that the
abbreviated description is, in part at least, a theory based on
epistemological considerations--belief, knowledge--and these have
nothing to do with the truth of modal sentences in which descrip­
tions occur. To circumvent this problem we might incorporate the
epistemological aspects of the theory into the descriptions them­
selves. Thus, strictly speaking, 'Homer' is short for 'the person
believed to be the author of the *Iliad*'. With this modification the
abbreviated description may be substituted for the name in modal
contexts. For example, it is true and compatible with our beliefs
that Homer might not have been an author since, by definition, this
means that the person believed to be the author of the *Iliad* might
not have been an author. On the other hand, it is necessarily true
that Homer is the person believed to have authored the *Iliad* but not
necessarily true that Homer is the author of the *Iliad*. It must be
granted that this device is somewhat artificial, but that does not
appear to prevent it from succeeding as an extension of the stronger
version of the abbreviated description theory to modal contexts.
If I am right in thinking that both Russell's and Kripke's theories
are best viewed as useful technical devices--or regimentations, to
borrow Quine's term--then artificiality is hardly objectionable
per se.


42. Ibid.

43. Ibid.

44. Kripke, IN, p. 78. It would be a short step to go from this remark
of Kripke's to the view that in our counterexample we are using the
word 'Nixon' as if it were a name in some other language than our
own. However, this sort of argument raises questions similar to
those raised above in discussion of Putnam's account of rigid
designation. See above pp.
CHAPTER VI
CONTINGENCY, NECESSITY AND RIGID DESIGNATION

The Contingent A Priori?

A view that is closely connected to the doctrine of rigid designation—indeed, a view that appears to follow from it—is Kripke's belief in the existence of contingent a priori and necessary a posteriori statements. He acknowledges that in holding this view he is going against recent empiricist thought which has held, for one reason or another, that the notions of a priority and necessity are coextensive. Kripke argues that the two notions must be sharply distinguished. He points out that on cognitive grounds the concept of a priority, as he calls it, is a concept of epistemology; it has to do with how we are able to know the truth of a statement: "a priori truths are those which can be known independently of any experience." In contrast, the concept of necessity, in at least one sense of its usage, is a concept of metaphysics; it has to do with the question of whether things might have been otherwise than they are:

We ask whether something might have been true, or might have been false . . . If [something] is true, might it have been otherwise? Is it possible that, in this respect, the world should have been different from the way it is? If the answer is 'no', then this fact about the world is a necessary one. If the answer is 'yes', then this fact about the world is a contingent one. This in and of itself has nothing to do with anyone's knowledge of anything.

Thus, as Kripke views it, necessity is a matter of truth in all possible worlds and this, according to Kripke, has nothing to do with knowledge. He concludes that in view of these obvious cognitive differences between
the two notions, "[i]f they are coextensive, it takes some philosophical argument to establish it."4

Kripke's argument for the view that necessity and a priori are not coextensive takes the form of several purported examples of contingent a priori and necessary a posteriori statements. One example suggested by Kripke concerns Leverrier's hypothesis that Neptune was the cause of certain perturbations in the orbits of the other planets. He asks us to assume that Leverrier gave the name 'Neptune' to the planet before it had been positively identified:

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. At that time he was unable to see the planet even through a telescope. At this state, an a priori material equivalence held between the statements 'Neptune exists' and 'Some one planet perturbing the orbit of such and such other planets exists in such and such a position', and also such statements as 'if such and such perturbations are caused by a planet, they are caused by Neptune' had the status of a priori truths. Nevertheless, they are not necessary truths, since 'Neptune' was introduced as a name rigidly designating a certain planet. Leverrier could well have believed that if Neptune had been knocked off its course one million years earlier, it would cause no such perturbations and even that some other object might have caused the perturbations in its place.5

As this example shows, Kripke's believe in the contingent a priori hinges on his thesis that names are rigid designators. The key claim in this example is that 'Neptune' "was introduced as a rigid designator."6

One problem with this sort of example is that there appears to be no way, even in principle, of distinguishing whether 'Neptune' was introduced as a rigid designator or as an abbreviation for 'the planet that causes the perturbations in the orbits of the other planets.' As Donnellan has noted, Kripke presents the following argument for the claim
that 'Neptune' was introduced as a rigid designator. Kripke argues that Leverrier might have believed (1),

(1) Neptune might have existed and not been the cause of the perturbations, when he introduced the term. But, as Donnellan points out, this argument is inconclusive because "anyone bent on maintaining that . . . the name was introduced as an abbreviation can escape this argument, . . . as Dummet does, by pointing to scope differences." 9

To see this consider the following example which is suggested by Donnellan:

(2) It might have been the case that the cause of the perturbations did not cause the perturbations.

(3) The cause of the perturbations might not have caused the perturbations.

The former sentence is clearly false; but the latter sentence is both true and compatible with those of Leverrier's beliefs—viz. (1) above—that Kripke regards as crucial. The only difference between (2) and (3) resides in the scope of the modal operator, and by clarifying these scope differences the Russellian description theorist can easily account for Leverrier's belief in (1) without abandoning the thesis that 'Neptune' is an abbreviation for 'the cause of the perturbations'. 10 Thus, as Donnellan suggests, we have something of a standoff here: We can account for the speaker's beliefs and language behavior on either theory. Therefore, to reiterate a familiar theme, no argument based on such evidence can decide whether a name has been introduced (or is used) as an abbreviated description or as a rigid designator. As Donnellan puts it, "in the absence of stipulation [one way or the other] . . . it would be indeterminate whether a name introduced by means of a description is a
rigid designator or an abbreviation, so long as the name continues to
be pegged to the description."\(^{11}\)

Although Kripke's argument does not show that 'Neptune' was actually
introduced as a rigid designator, Donnellan argues that it does suffice
to show the "theoretical possibility" that a rigid designator may be
introduced by means of a description.\(^{12}\) According to Donnellan, we may
stipulate that a name is a rigid designator or, on the other hand, that
it is an abbreviation: If it is specifically intended and stipulated
that this introduction of a name is to be the introduction of a rigid
designator, I see no theoretical reason to suppose that the stipulation
cannot accomplish what is intended."\(^{13}\)

However, despite Donnellan's great faith in the efficacy of such a
stipulation, I should like to argue that there are serious theoretical
obstacles involved in such a procedure. First, as Donnellan notes, to
stipulate that 'Neptune' is a rigid designator is tantamount to taking
a sentence like (4),

(4) Neptune is the planet that causes the perturbations,
as contingently true. (4) is contingent on Kripke's theory because there
is a world in which 'Neptune' and 'the cause of the perturbations'
designate different objects. Donnellan acknowledges this by putting the
stipulation itself in the following terms:

If we want to introduce the name 'N' [as a rigid designator]
by means of the description 'the φ' then the formula we would
use would be: Provided that the φ exists, let 'N is the φ'
express a contingent truth.\(^{14}\)

As Donnellan himself is aware, viewing matters in this way leads to
problems; for it looks as if we here gain knowledge of a contingent
fact by a mere act of stipulation.\(^{15}\) To circumvent this problem Donnellan
does not question the idea that a name may be stipulated to be a rigid designator; rather, he denies that in such cases we gain knowledge of the world. The key to his argument is a distinction between "knowing that a certain sentence expresses a truth and knowing the truth of what is expressed by the sentence."\(^{16}\) This distinction is illustrated as follows:

A person could know that a certain sentence expresses a truth if he has been told by an unimpeachable source that it does. Thus I could learn from a German speaking friend that a certain German sentence expresses something true, but if I do not speak German and do not know what the sentence means I would say that I do not know the truth of what the sentence expresses.\(^{17}\)

Thus, according to Donnellan, we can distinguish between 'knowing that "p" is true' and 'knowing that p'.

On the basis of this distinction Donnellan concludes, with some additional argument which we shall consider presently, that knowledge that the sentence 'Neptune is the cause of the perturbations' expresses a truth does not count as a priori knowledge of a contingent fact because it is mere knowledge that a sentence is true and not knowledge about the world.

However, there is a problem here; for to claim on the one hand that Leverrier lacked knowledge of the world—a claim which we shall examine below—appears to conflict with the very notion of designation which, after all, is a relation between an expression and an object of the world. In short, not only does Donnellan's distinction conflict with the notion of rigid designation, it appears to conflict with the notion of designation itself. To see this let us consider an example that Donnellan has devised.\(^{18}\) Suppose that a detective who is investigating
an apparent murder proposes that 'Vladimir' be the name of the murderer who, in fact, is the murderer. If Jones is the murderer, then this sentence corresponds to that state of affairs in which Jones committed the murder. But, as Donnellan argues:

... someone [other than Jones] might have been the murderer; in other possible worlds it is Smith or Robinson. We can imagine everything being the same except that someone else is the murderer (and whatever differences are entailed by that). Then a different state of affairs would correspond to the sentence ['Vladimir is the murderer'] and it would express the existence of a different state of affairs. We could also put this by saying that in different possible worlds 'Vladimir', as the name is introduced by stipulation in those worlds, would turn out to rigidly designate someone else. 19

We can agree with Donnellan that in proposing the name 'Vladimir' the detective does not obtain knowledge of which of these possible states of affairs obtains; assuming that Jones is the murderer, the detective does not obtain knowledge that Jones is the murderer. 20 This is the element of truth in Donnellan's argument. So the detective does not know whether Jones or Smith or Robinson is the referent of 'Vladimir'. But if 'Vladimir' can designate different objects in different possible worlds—as Donnellan assumes—how can it be claimed to be a rigid designator? If it can designate different objects in different possible worlds, then, by definition, it is a non-rigid designator. Moreover, as here introduced, the name 'Vladimir' designates whoever is the murderer and not some identifiable person; as such however, it possess precisely the same criterion of identification as 'the murderer' or, what amounts to the same thing, its referent is determined by precisely the same identity conditions as the referent of 'the murderer'. Because of this,
it is difficult to see how one expression can be said to designate rigidiy and the other non-rigidly. Therefore, Donnellan's claim that 'Vladimir' is here introduced as a rigid designator seems highly arbitrary and totally without justification.

In response to this it might be argued that 'Vladimir' does not designate different persons in different possible worlds; rather, it designates one and the same person in all possible worlds but we just do not know who that person is. But this type of response merely involves us in a more pressing problem—viz. how can it be claimed that an expression designates the same object in all possible worlds when we cannot even identify its referent in the actual world? Does it make sense to talk of an object being the same if we can't identify the object in the first place? A key contention of Kripke's rigid designation thesis is the claim that the referents of rigid designators are objects "which we have, and can identify, in the actual world." Now, it must be admitted that identification is a matter of degree. But surely, it would be stretching our credulity to maintain that we "have" the referent of 'Vladimir' and can identify it in the actual world for, as the example is constructed, it is not even established that there is a murderer. The same can be said for Kripke's Neptune example. Indeed, the fact that such names as 'Vladimir' and 'Neptune', when introduced, may designate different objects, or no objects at all, appears to be pretty firm evidence that here the names are really introduced as abbreviations and not as rigid designators.

Moreover, to return to the point made at the outset of this argument, if it is the case that Leverrier or the detective lacked knowledge of the
world when they introduced such names, this appears to imply that such names are not designators at all; for designation is a relation between an expression and an object of the world. Again, this suggests rather forcefully that such names are abbreviated descriptions which, by Russell's theory at least, would be analyzed in terms of the quantifier-variable notation and therefore not construed as designators at all.

What then of Donnellan's claim that the above "stipulations have not given rise to any knowledge (other than of linguistic matters) [and] so not to any knowledge a priori"?23 In the first place, it should be noted that Donnellan's notion of 'knowing that a sentence expresses a truth' has virtually nothing to do with the concept of a priori knowledge. For, according to Donnellan's argument, 'knowing that "p" is true' boils down to 'knowing that "p" is believed true by some authority'. But this sort of knowledge-by-authority simply postpones the question of whether 'p' can be known independently of experience which is the question we must ask if we are interested in determining whether 'p' is knowable a priori. Moreover, in the examples we are considering, it does not even appear that Donnellan's notion would be applicable; for surely, unlike his own illustration of 'knowing that "p" is true'—whereby one can know the truth of a German sentence without knowing what it means—we must assume that Leverrier knew what the sentence 'Neptune is the cause of the perturbations' meant when he introduced the name 'Neptune'. And it appears reasonable to assume that knowing what a sentence means amounts to knowledge of the world.24 Of course, if 'Neptune' functions here as a disguised description, then the sentence appears to mean merely that 'the cause of the perturbations is the cause of the perturbations'.
This sentence is trivially true, but then so are most—some would say all—sentences whose truth we know a priori. As Frege might have viewed it, Leverrier's stipulation did not constitute a very valuable extension of his knowledge of the world, but then neither does our knowledge that a = a. Yet, in this latter case, we feel no temptation to claim that in knowing that a = a we lack genuine a priori knowledge.

One further point seems worth mentioning. It might reasonably be argued that such stipulations lead us only to knowledge of the way signs are used, and if we exclude knowledge of signs from proper knowledge—as Frege himself seems to have done—we might be inclined to argue that such stipulations do not lead to genuine knowledge of the world. But, in response to this, it should be remembered that signs are part of the world, and therefore it appears entirely unfounded to exclude knowledge of signs from genuine knowledge. Thus, no matter how you slice it, it does not appear reasonable to claim that such stipulations do not lead to genuine a priori knowledge: there is nothing inherent in the concept of a priori knowledge that limits its application to knowledge of extra-linguistic entities.

Thus far, we have been concerned with alleged examples of the contingent a priori and with Donnellan's attempt to show that such sentences are contingent but not a priori. We have seen that the alleged contingency of such sentences follows directly from the claim that the names involved are purported to be rigid and the descriptions non-rigid designators. Such examples prove little since if we construe the names as abbreviated descriptions, the sentences turn out to be analytic. Moreover, we have adduced reasons for rejecting Donnellan's claim that
such sentences are not genuine a priori truths. Thus, if we view the names as abbreviated descriptions—which in the examples considered seems the more plausible view—the sentences turn out to be necessary a priori: They become tautologies when the name is replaced by its associated description; e.g., 'N is the φ' becomes 'the φ is the φ' when the name 'N' is replaced by the description, 'the φ', which it abbreviates. Therefore, Kripke's claim about the contingent a priori inspires little confidence.

The Necessary A Posteriori?

Kripke's doctrine of rigid designation also leads him to claim that certain sentences are necessarily true but knowable only a posteriori. It is to such examples that we now turn.

Kripke's examples include the following sentences:

(5) This table (pointing) is composed of molecules.
(6) Gold has atomic number 79.
(7) Water is H₂O.

In each case above the underlined expression is said to be a rigid designator, and in each case Kripke claims that the necessity involved is metaphysical necessity—i.e., truth in all possible worlds. So for each case we cannot say truly that, e.g., 'Gold might not have had atomic number 79'. Moreover, since in each case the fact expressed by the sentence is a fact that was discovered (or discoverable) rather than a matter of mere stipulation, we know these necessary truths a posteriori. Finally, in each case an alternative way of expressing the discovery is in terms of essential attributes. Thus we can say that gold is necessarily
the substance with atomic number 79, or that having atomic number 79 is an essential attribute of gold.

Kripke's gold example is particularly interesting because it is parallel, in certain respects, to the Neptune example. Consider the two sentences involved side by side:

(8) Gold is the substance with atomic number 79.
(9) Neptune is the planet that causes perturbations . . .

In each case the underlined expression is the rigid designator and in each case the rigid designator is pegged to a definite description. But, according to Kripke, it is only in the case of (9) that one can truly say '_____ might not have been the $\phi$' where the blank is filled by the rigid designator. Well, one can say 'Gold might not have had atomic number 79' but Kripke means that it would be false to say that. This is so, presumably, because one cannot imagine a possible world in which (8) is false. But this appears arbitrary, for it is quite imaginable to me that the stuff that we call 'gold' might not have atomic number 79. To see this let us consider the following situation.

The truth of (8) depends on scientific theory. Currently, scientists hold true a certain body of theory, say the set \{T_1, T_2, \ldots, T_n\}, from which it follows that gold has atomic number 79. But surely it is possible that current scientific theory is mistaken; it's happened before that our scientific theories were overthrown in favor of better theories. Let us call such a world in which current scientific theory is false $W_b$. Then suppose that in $W_b$ the body of theory, \{T_1', T_2', \ldots, T_n'\}, is true and that it follows that in $W_b$ gold does not have atomic number 79. We are here talking about the very same stuff that scientists
in the actual world call 'gold'—viz. the stuff which is found in certain mountains, used as a medium of exchange, and so on. Now this appears to me to be a situation which raises serious difficulties for Kripke's claim that gold is necessarily that stuff with atomic number 79, for it is a possible world in which (8) is false. Kripke disagrees however. He argues that a world in which we found, say, iron pyrites in those mountains where in the real world we find gold, "would not be a case in which possibly gold might not have been an element [of atomic number 79], nor can there be such a case (except in the epistemic sense of 'possible')."\^31

But this really isn't playing very fair on Kripke's part. I imagine a possible world in which (8) turns out false and then I'm told it doesn't count because it is only an epistemically possible world. Not only is there an element of foul play here, there appears also to be an element of inconsistency on Kripke's part. For why aren't the worlds in which possibly Neptune is not the cause of the perturbations merely epistemically possible? Surely, if I choose, I can describe our world \(W_b\) a bit more fully in the following way: In \(W_b\) violent reactions on the sun have affected the gravitational field of the solar system in such a way that Neptune no longer causes any perturbations in the orbits of the other planets. So I've here imagined a possible world in which each of (8) and (9) are false, and therefore all must be contingent truths if we adopt Kripke's standards. But according to Kripke, the contingency involved in (8) is epistemic while that involved in (9) is supposed to be metaphysical. Yet it is one and the same imagined world that makes each of (8) and (9) false. If this is not inconsistent, then what is wrong with our possible world \(W_b\)?
It is not difficult to see why Kripke thinks it impossible for there to be a world in which gold did not have atomic number 79. He builds the impossibility right into his argument. In that example he asks us to suppose that "scientists have investigated the nature of gold and have found that it is part of the very nature of this substance, so to speak, that it have atomic number 79." He then later argues that "given that gold is this element, any other substance, even though it looks like gold and is found in the very places where we in fact find gold, would not be gold. It would be some other substance which was a counterfeit for gold." He also claims that "we wouldn't say that [this counterfeit substance] is gold ... One should not say that it would still be gold in this possible world, though gold would then lack the atomic number 79. It would be some other element, some other substance." As these remarks reveal, what Kripke is doing here is defining 'gold' in terms of the description 'the substance which has atomic number 79'. Having atomic number 79 is a necessary condition for being gold and anything which fails to satisfy this condition is not gold. Thus, what we have here is not only a case where the referent of a rigid designator is picked out by a description, but a case where the rigid designator is analytically tied to the description.

Moreover, notice the disparity between the gold example and the Neptune example. In the gold example the referent of the rigid designator 'gold' is picked out by the description 'the substance with atomic number 79', and we are told that 'Gold has atomic number 79' is metaphysically necessary. In the Neptune example the referent of 'Neptune' is picked out by the description 'the planet which causes certain perturbations', but the sentence 'Neptune is the planet which causes certain
perturbations' is said to be known a priori. However, there is no appreciable difference between the two cases.

We see then that Kripke's employment of the notions of a priority and necessity are quite arbitrary and not terribly consistent. One who thought these undesirable qualities for any theory to possess might conclude that Kripke's theory is false. A more reasonable conclusion, however, is that Kripke has simply failed to give clear meaning to a notion of metaphysical necessity that is independent of epistemological considerations. His notion of metaphysical necessity is just analyticity in disguise.

To see this consider the gold example again. Semantically, one could justify the contention that (8), 'Gold is the substance with atomic number 79' is necessarily true, if one could show that there is no possible world in which (8) is false. One can do this, however, only if we place restrictions on the possible worlds so as to rule out examples like that constructed above. One such condition (which works) is that the possible world must be consistent with what we know or with what we hold true. Then, since we know that gold has atomic number 79, there will be no possible world, in this new sense of possible, in which (8) is false. Therefore, (8) is necessarily true in this new sense. Colloquially, we might characterize this sense of necessity as 'necessary given what we already know'. '5+5 = 10' is necessary in this sense; so are true sentences about the past--e.g., 'Nixon was President in 1970'. Sentences about the future, on the other hand, are merely contingent in this new sense of 'possible'.
This sense of necessity is transparently tied up with epistemological considerations. One such epistemological assumption is that the word 'gold' stands for that substance which has atomic number 79. This has nothing to do, ostensibly, with whether or not language users employ the word 'gold' to refer to such a substance. If they do we would say they speak correctly; if not then they speak incorrectly. But, again, isn't this just analyticity all over again? Surely, on intuitive grounds, as Mill pointed out long ago, there is no impossibility involved in conceiving of a thing which has all the superficial qualities of gold but lacks its atomic structure; the impossibility, if there is any, stems from the way we have defined 'gold': we just wouldn't call (correctly, that is) such a substance 'gold'. Kripke sees all this but still insists on viewing the necessity involved as metaphysical necessity "in the strictest possible sense."\(^3\)

We might agree with this if Kripke could provide a semantic analysis of 'Gold is necessarily the substance with atomic number 79' that did not involve epistemological considerations. But this does not appear very likely. Any such analysis must prohibit the possibility that (8) is false in some possible world. There appear to be two ways to do this: One is to claim that if (8) is false in some possible world, it is only because in that world the term 'gold' is being used incorrectly—\(\text{we just shouldn't say that such a substance is gold. But this is an epistemological argument that presupposes a certain definition of 'gold'.} \) The other way is to claim that any such possible world (in which (8) is false) is merely epistemically possible. But this is ambiguous. Surely, we know that gold has atomic number 79 and we might not have
known this. That's one sort of epistemic possibility, I suppose. Another is that what we claim to know might be mistaken. We say we know that gold has atomic number 79, but what if it doesn't? Then our alleged knowledge is mistaken which is another way of saying that gold might not have atomic number 79. This is not 'epistemically possible' except in the trivial sense that knowledge of some sentence always implies the truth of that sentence. But it is in this latter sense that our possible world $W_b$ takes (8) to be false: We have not assumed that in $W_b$ the scientists might not know such and such; we've assumed that they know (or believe they know) our present theories to be false. Therefore, it is difficult to see how Kripke's invoking of 'epistemically possible' can save his thesis from a possible world in which (8) is false. To conclude, then, there appears to be no way for Kripke to provide a semantic interpretation of 'Gold is necessarily atomic number 79' that yields the value true for this sentence (in the actual world) and does not invoke non-trivial epistemic considerations. If one insists on the necessity of 'Gold has atomic number 79', it appears that one does so only because one has defined 'gold' in terms of its atomic number. This is just analyticity.

Of course, in the Neptune example Kripke claims that (9), 'Neptune is the cause of perturbations', is contingent on the basis that one can truly say 'Neptune might not have been the cause of the perturbations'. But what is it about the expressions 'Neptune' and 'the cause of the perturbations' that allow one to say this but do not allow one to say 'Neptune might not have been Neptune'? It seems to me that the crucial difference is really that difference between sentences of the form 'a=b'
and 'a=a' which exercised Frege's mind. We know that 'a=a' is true a priori. It is also logically true as an instance of the general logical truth that everything is self-identical. We do not know a priori that 'a=b' is true unless we know that 'a' and 'b' designate the same object. This we know in the Neptune example by stipulation and therefore 'a=b' is known a priori in that case. Nevertheless, it is not logically true because it would not remain true under any and all re-interpretations of the non-logical symbols a and b. However, following Carnap, we can account for the analyticity (i.e., the necessity) of 'Neptune is the cause of the perturbations' if we adopt it as a "meaning postulate". In a way this is no different from what Kripke did in the gold example: in both cases it is required that any possible world in which 'a=b' (where 'a' and 'b' are replaceable by the appropriate terms from the gold and Neptune examples) is necessary since we know (or hold true) that 'a' and 'b' designate the same object. In this way Kripke's Neptune example can be brought into parallel with his gold example, but only at the expense of dropping his contention that 'Neptune is the cause of the perturbations' is a contingent truth.

The above considerations appear to account well for our intuition that, as Kripke has set up his examples, both (8) and (9) above are analytically true. A possible objection suggests itself at this point. In (9), the Neptune example, the connection between the rigid designator and the description is stipulative but in (8), the gold example, it is discovered. Therefore, (8) should be an a posteriori truth, not analytic a priori. In response to this it should be noted that the fact that (8) was discovered originally (and not stipulated) drops out of consideration
in Kripke's argument. For his claim that gold is by its very nature that stuff with atomic number 79 is just a standard way of defining 'gold' in terms of some property. It is the fact that Kripke allows no possible case in which gold lacks this property that belies the definitional nature of (8). Therefore, (8) is analytic in precisely the same sense as Katz's above claim that 'cats are animals' is analytic.
NOTES--CHAPTER VI


3. Ibid., p. 261.

4. Kripke, IN, p. 85.


6. Ibid.


8. See Kripke, NN, pp. 347-348.


10. See above pp. 100 ff. for a discussion of scope ambiguities.

11. Donnellan, p. 16. It is also worth noting that on an abbreviated description theory we would have grounds for regarding 'Neptune is the planet that causes perturbations' as necessarily true, since it becomes a tautology when we substitute the description for the name. Thus, again, the dispute boils down to what appear to be unsolvable problems: Is 'Neptune is the planet that causes perturbations' contingent or necessary? The answer appears to hinge on which semantics we choose, Kripke's or Russell's, and does not admit of a non-question begging answer.

12. Ibid.

13. Ibid.


15. This difficulty is pointed out by Dummett, op. cit., p. 121.

17. Ibid.

18. Ibid., p. 20.

19. Ibid.

20. We can agree that the detective does not attain knowledge of Jones, but it doesn't follow that he thereby fails to possess a priori knowledge. For the sake of argument assume that 'Vladimir' is just an abbreviation for 'the murderer'. In that case 'Vladimir is the murderer' is synonymous with 'The murderer is the murderer'. It becomes a tautology—a necessary truth—when we substitute the description for the name that abbreviates it. Now this constitutes genuine a priori knowledge because the detective needn't appeal to experience to know that 'Vladimir is the murderer' is true. So Donnellan appears to be confusing two distinct questions here: knowledge of a distinct individual who has been identified—Donnellan calls this knowledge de re (p. 21)—and a priori knowledge. The two notions do not appear to be the same. It might even be granted that a priori knowledge is de dicto knowledge—i.e., knowledge of the truth of 'Vladimir is the murderer'. But so long as the truth of this sentence can be known independently of any experience, I see nothing amiss with this claim. No one ever said that a priori knowledge must be knowledge de re in Donnellan's sense.

21. This point is related, in certain respects, to the problem of transworld identity. (Discussed in the next note.) Kripke maintains that the notion of rigid designation may be conceived independently of epistemological considerations. But such questions as we have been considering cast considerable doubt on this claim. Moreover, this is not particularly surprising in light of the fact that naming, reference and designation, insofar as these notions apply to ordinary discourse and not merely to formal semantics, involve us in epistemological questions.

22. Kripke, NN, p. 273. Kripke employs this argument to circumvent the widely discussed problem of transworld identity. The problem is this: Kripke's notion of rigid designation depends on the claim that the same object exists in different possible worlds. Several authors have pointed out that it is meaningless to talk of an object being the same in another possible world if we cannot provide identity conditions for that object. To take an example, consider the name 'Nixon' and that possible world in which Nixon lost the 1968 Presidential election. According to Kripke, we are talking about Nixon here, the same person who was President in 1970, had a dog named "Checkers", and so on. According to Kripke, to even think that there is a problem of identifying Nixon in some possible world stems from the false picture "that a particular is nothing but a 'bundle of qualities". (NN, p. 272.) According to this false picture, in talking about a different possible world in which Nixon was not President we are really talking about a different bundle of qualities than the Nixon of the real world. Moreover, Kripke argues
that the problem of transworld identity is based on a false picture of 'possible worlds'. He argues that possible worlds are stipulated, not discovered—as if they were distant planets—so the problem of identifying objects in possible worlds is really a pseudo problem. His reasoning is that identification is an epistemological problem and "[p]ossible worlds are not something to which an epistemological question like this applies." (NN, pp. 81-82.)

There is some merit to Kripke's argument and this should not be denied. It does seem to be rather idle to ask of a possible world 'How do we find out if the person you're talking about is really Nixon?' Nevertheless, there is still a problem here and one whose significance Kripke's argument appears to miss. It is this: If possible worlds are stipulative, as Kripke insists, and if we are not compelled to provide even minimal identity conditions for the objects of such worlds, then it is equally idle to argue over the necessity or contingency of English sentences. In other words, to eschew talk of identity conditions appears to leave us with a situation whereby anything goes as regards the necessity or contingency of certain sentences. Take, for example, the sentence 'Nixon might have been a cucumber'. Is it true? Is 'Nixon is a cucumber' true in some possible world? We want to say, of course, that it is not true, but the only way to support this claim is to claim that objects have certain essential properties such that it is impossible for Nixon to have been a cucumber. This Kripke does. He notes that "some of [Nixon's] properties may be essential." (NN, p. 272) But to acknowledge this is to acknowledge that there is a problem of transworld identity here, for to claim that an object has certain essential properties is just to claim that we do possess certain conditions by which we identify that object. Thus, despite what Kripke says about this problem, we are enmeshed in epistemological questions when we talk about an object being the same in different possible worlds. Moreover, until identity criteria are established, we possess no basis for the distinction between necessary and contingent sentences; all such debate becomes mere idle chatter unless we can provide grounds for ruling out possible worlds in which Nixon is a cucumber. To my mind, such grounds would be epistemological in nature.


24. Knowledge of meanings, after all, is extra-linguistic knowledge.

25. See above, note 20.

26. See Gottlob Frege, 'On Sense and Reference', in Peter Geach and Max Black (eds.), Translations from the Philosophical Writings of Gottlob Frege (Oxford: Basil Blackwell, 1970), pp. 56-57: Frege argues that if a sentence like 'a=b' was about the signs 'a' and 'b', it "would no longer refer to the subject matter, but only to its mode of designation; we would express no proper knowledge by its means."

27. See note 11 above.

28. Here I do not assume, with Russell, that descriptions like 'the \( \phi \)' are analyzed in terms of '(Ex)...', in which case the sentence 'the \( \phi \) is the \( \phi \)' would be a synthetic statement asserting existence and not a tautology.

29. See Kripke, NN, pp. 314-315.

30. As we noted above, Kripke attempts to disassociate metaphysical necessity—truth in all possible worlds—from epistemological considerations. We have learned to become increasingly skeptical as our argument proceeds. See below pp. 151 ff.

31. Kripke, NN, p. 320.

32. Ibid., p. 319.

33. Ibid., p. 320.

34. Ibid.

35. Ibid.


CHAPTER VII

THE ATTACK ON THE PRINCIPLE OF IDENTIFYING DESCRIPTIONS

Donnellan's Argument

In the two preceding chapters we have considered Kripke's rigid designation theory which, as we noted, was designed to refute the claim that the meaning of a name is just a definite description; that is, Kripke's rigid designation theory of names is conceived as an alternative to the principle of identifying descriptions (PID) taken as a theory of meaning. We have found Kripke's and Donnellan's arguments to be lacking in persuasive force and therefore, for at least some proper names—e.g., 'Homer', 'St. Anne', 'Neptune' as it was portrayed in Kripke's example, and many others—we have as yet no reason to believe that Russell's theory is false. Of course, it is doubtful that Russell's theory can be extended to all names—e.g., to personal names like 'Richard Nixon', place names, etc.—except perhaps in its weaker (and more artificial) version whereby such names are short for descriptions of the form 'the person named "N"'. In this chapter we shall take up criticisms of the principle of identifying descriptions interpreted as a theory of reference rather than as a theory of meaning. The key difference between the two interpretations is that on the former it is not claimed that the associated descriptions are synonymous with the name—or give the definition of the name—but only that they pick out the intended referent of the name.¹ We save for the next chapter the discussion of Donnellan's and Kripke's alternatives to the identifying description theory of reference.
In 'Speaking of Nothing' Donnellan asserts boldly that "the principle of identifying descriptions is false." He takes that principle—PID, for short—to be that names "by one mechanism or another . . . are surrogates for descriptions"; and he takes this as applying both to Russell's disguised description theory and to Searle's view that names are correlated with descriptions. The view that Donnellan hopes to refute is that names are correlated with a set of descriptions and what one is saying in, say, a simple subject-predicate statement employing a proper name is that whatever best fits these descriptions has whatever property is designated by the predicate.

This formulation of the PID is somewhat misleading for it does not appear to allow the possibility that some of the descriptions associated with names may be mistakenly believed to apply to the intended referent without thereby hindering the ability to refer to that referent via the name. Donnellan's formulation suggests that in order to refer to, say, Socrates by using the name 'Socrates', Socrates must "fit" the associated descriptions. If he doesn't, we have failed to refer to a unique individual. Actually however, the PID presents a slightly different picture. It is quite willing to allow that some of the descriptions believed true of Socrates are actually false of Socrates, and it requires only that some of the descriptions must be true of Socrates. As Wittgenstein put it, if some of our props fail then we can rely on others. The central point of the PID is that successful reference involving proper names can only occur against a background of shared beliefs about the intended referent. Therefore, in order to refute the PID, Donnellan must show that successful reference involving names can
occur in situations where all the associated descriptions are presumed to be false.

Donnellan's argument runs as follows:

... supposing you could obtain from me a set of descriptions of who it is that I believe myself to refer to when I say 'Socrates was snub-nosed'—perhaps such things as 'the mentor of Plato', 'the inventor of the "Socratic method"', 'the philosopher who drank the hemlock', and so forth—it is theoretically possible that I am referring to something about which no substantial number of these descriptions is true or that although there is something that fits these descriptions to whatever extent is required by the particular version of the principle, that is not in fact the referent of the name I used.\(^7\)

In and of itself, the "theoretical possibility" to which Donnellan here refers does not constitute a refutation of the PID. For, according to this principle, if it turns out that all, or some substantial number, of the identifying descriptions are mistaken, it would follow simply that 'Socrates' does not refer to a uniquely identified individual; this is to say that if all the descriptions associated with the name turn out false, then 'Socrates' could not be used to refer to a unique individual. The proof of the pudding is that we should not then be able to say to whom we are referring when we use the name 'Socrates'. Donnellan's point is that 'Socrates' could still be used to refer to a unique individual in spite of the fact that all (or a substantial number) of its associated descriptions are false. According to Donnellan, this is possible because "successful reference will occur when there is a historically correct explanation of who it is that the speaker intended to predicate something of. That individual will then be the referent..."\(^8\)

This argument constitutes the central claim of Donnellan's 'Historical Explanation Theory of Proper Names' (HET), a theory which in its general outlines at least, is very similar to Kripke's 'Causal Theory'
(CT) of names. We shall discuss the particulars of each theory below, but
since one of Kripke's and Donnellan's main reasons for proposing their
theories is their belief that the PID is false, we shall begin by looking
at the negative side of their arguments—viz. their criticism of the PID.

Our task then is to determine in precisely what sense it could be
'theoretically possible' to use a name to refer to an individual about
whom all our presumed information is misinformation. Donnellan's theory
provides the following answer:

... if we take the set of identifying descriptions the speaker
could give were we to ask him to whom he was referring, 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... it is
clearly imaginable that our omniscient observer sees that
while the author of the dialogues did intend one of the charac-
ters to be taken as a portrayal of a real person, he modestly
attributed to him a method that was his own brain child. And
in general, it would be possible to have the historical con-
nection with no end to mistaken descriptions in the head of the
speaker... 

What then are we to make of Donnellan's claim?

First, to forestall any confusion over a rather non-essential point,
let us drop the notion that it is the descriptions in "the head of the
speaker" that are relevant. Individual speakers are members of a
linguistic community and it is in virtue of the shared beliefs among
members of the community that successful reference is possible at all.
Although one particular speaker may not be able to provide a correct
identification of the referent of 'Socrates', this does not detract from
his ability to use that name. The reason is that other speakers can
provide such an identification, and this is all that is required to
provide the descriptive backing regarded as necessary by proponents of
the PID.
Kripke presents a number of examples which attempt to exploit this subjective interpretation of the PID. For example, he argues that a speaker may refer to Einstein by employing the name 'Einstein' even if he can't produce a very good explanation—a uniquely identifying, non-circular description—of who Einstein is. Kripke believes this sort of argument refutes the PID. Another example he employs goes as follows:

Consider Richard Feynman, to whom many of us are able to refer. He is a leading contemporary theoretical physicist. Everyone here (I'm sure!), can state the contents of one of Feynman's theories so as to differentiate him from Gell-Mann. However, the man in the street, not possessing these abilities, may still use the name 'Feynman'. When asked he will say: well he's a physicist or something. He may not think that this picks out anyone uniquely. I still think he uses the name 'Feynman' as a name for Feynman. It is true that the man in the street would be using the name 'Feynman' as a name for Feynman. But in such cases his ability to refer successfully and to be understood as referring to a unique individual is parasitic on another's ability to identify Feynman. As Dummett points out, we frequently use names for which we can give no very good explanation. An extreme case would be situations like when one relays to one's roommate that 'Mr. Cunningham telephoned today'. In such cases we act as a mere "recording apparatus"; "but there is no sharp line between such a case and a full fledged mastery of a name: a whole series of transitional cases stretch from one to the other." When we employ a proper name for which we cannot provide a unique identification of its referent, "we are exploiting the fact, known to us, that the word we use is part of a common language [and that the referent] could be more precisely identified ... by our hearers." A similar argument would apply to the use of 'Feynman' by the man in the street.
We might add that proper names are not merely pulled out of the air by the man in the street and thereby used correctly. They are usually passed on from speaker to speaker, and, generally, this is done together with at least a partial identification—in terms of descriptive information—of the bearer of the name. Even if the man in the street obtains the name 'Feynman' by overhearing it used in a conversation between third parties and is not introduced to the name directly, he cannot help but pick up some information about Feynman. (This assumes that the 'conversation' does not consist of someone just yelling the name 'Feynman' but rather in the name being used in certain sentences.) Therefore, it is doubtful that one could acquire the use of a name without acquiring some information about its referent. But even if one could acquire a name without also acquiring some information about its referent, this would not refute the PID so long as we suppose that other speakers within the community possess the ability to identify the bearer of the name.

We see, then, that in order to refute the PID Donnellan must show that successful use of names can occur when the beliefs operative within the language community are mistaken, not merely when the beliefs of some particular speaker are mistaken. Donnellan believes that this is 'theoretically possible'. That is, it is theoretically possible that when we use the name 'Socrates' we are thereby referring to an individual about whom all our present beliefs are false. This appears to me to be not even theoretically possible.

One problem with Donnellan's argument is that his appeal to an omniscient observer constitutes an illegitimate broadening of the
language community in which the name 'Socrates' has a use.\textsuperscript{16} The knowledge and beliefs of the omniscient observer go beyond the knowledge and beliefs of members of the language community. We have throughout characterized the PID as holding that successful reference can occur only against a background of shared beliefs. Part of this background consists of the identity conditions of the bearer of the name 'Socrates'; i.e., an explanation of who Socrates is. As we have said, it is quite possible that we could be mistaken in our beliefs, even in all our beliefs about Socrates; but in that case we would not be referring to a unique antecedently identifiable individual—at least not in any straightforward sense of 'referring'. What Donnellan claims, in effect, is that we could refer to an individual which we—i.e., the language community—have not identified in any way. How could we identify him since \textit{a fortiori} we know nothing about him, not even that such an individual exists? On intuitive grounds at least, Donnellan's 'theoretical possibility' does not strike me as very plausible.

What's worse, it leads to absurd conclusions. For example, it follows from Donnellan's argument that it is possible to refer to an individual that one cannot possibly have intended to refer to. The act of reference, as Donnellan is well aware, is an intentional act. In general, to refer to an individual I must \textit{intend} to refer to that individual; that is, I must believe, antecedently to my act of reference, that there is some unique individual, distinguishable from others in some way, and I must intend to refer to that individual. Similarly, to understand a unique reference, the audience must believe that I am intending to refer to a unique individual and must also be able to
identify the intended referent in some way or other. (Not necessarily in the same way that I would identify the referent.)

Of course, there are certain exceptions to this rule. For example, in one sense of 'refer' I could intend to refer to Smith but through some mistake or other I refer to Jones. Say, for instance, I see someone mowing Smith's lawn and I believe it to be Smith but it is actually his neighbor Jones who, unbeknownst to me, has lost a bet to Smith and is paying up by mowing Smith's lawn. Suppose then I remark, 'Look at Smith over there. That guy is positively fanatic about keeping his lawn trimmed; he mows it at least three times a week.' Now in such a situation we might wish to say that I intended to refer to Smith---after all, it is Smith who I consider to be fanatic---but have actually referred to Jones; i.e., I referred to Jones at least in that part of my remark when I said 'Look at Smith'. Even this example is not a clear case of referring to an individual to whom I did not intend to refer, for it could be argued that when I said 'Look at Smith' I meant to (and did) refer to the person mowing the lawn, but when I said 'That guy . . .' I intended to and did refer to Smith. So it is not a clear-cut case of referring to an individual to whom I did not intend to refer. But even if it were, it would not be in the same league with the type of case that Donnellan is urging. His is what we might call a case of radical failure of intention; not only do I refer to an individual to whom I do not intend to refer, but, as far as I know, there is no such individual. At least in the Smith case I know of two distinct individuals, Smith and Jones, although I do not believe Jones to be present. Or, if I don't know of Jones, I am or can be made aware of the possibility that someone else
besides Smith might be mowing Smith's lawn. The failure of intention which Donnellan's argument implies is far more radical and, indeed, appears to be impossible.

Donnellan's example involves a case of referring to an individual about whom I cannot possibly know anything at the time of reference. For, were Donnellan to assume that the speaker knows or believes that Socrates did not invent the Socratic method, etc., then the speaker could employ this information to distinguish the intended referent and the case would no longer constitute a counterexample to the PID. To see this let us assume that Donnellan's omniscient observer sees that the character Socrates was modelled after Plato's fishmonger uncle who was a virtuous person but philosophically naive and who died in a shipwreck off the coast of Athens in 401 B.C. Assume that this new information about the real Socrates is contained in a recently discovered diary written by Plato himself. What are we to say about referring uses of 'Socrates' both before and after the discovery of the diary? Did all such uses, both before and after, refer to the real Socrates? Or did those before the discovery refer to some individual about whom we believed certain things to be true? Certainly, one thing we should say here is that there is no clear-cut answer available independently of what intentions were actually involved in the actual speech context. What if, before the discovery, it was said 'Socrates was snub-nosed'? Is it the real Socrates or the person we believe to be Socrates that is said to be snub-nosed? No clear-cut answer emerges when the question is posed in this way. We could say just about anything we like about such an isolated sentence without offending our intuitions about the case.
Assume for example, that information in the diary includes a description of Plato's uncle and assume that his uncle had such a prominent nose that he was nicknamed 'the Snoz' or something. Then the sentence 'Socrates is snub-nosed' is false of Plato's uncle but true of the person we formerly believed to be Socrates. Before the discovery we could then say either of two things about the use of 'Socrates'. We could say it referred to the person believed to be Socrates and made a true prediction about him, or that it referred to Plato's uncle and made a false assertion about him. Neither explanation appears obviously counter-intuitive. Is there any way to decide which explanation is the correct one?

One problem with our example is that it attempts to draw conclusions from too isolated an utterance. Therefore, let's fill-in the context a bit more. Suppose the context was a classroom, a lecture about Socrates that occurred before the discovery of the diary. And suppose the lecturer said the following: 'Socrates was really the paradigmatic philosopher. He was extremely virtuous; his actions were always the consequence of his beliefs about what was right; he was devoted to the pursuit of truth and he pursued this aim at the expense of all other things, including, at times, his own safety. The portrayal of Socrates in Plato's dialogues is modelled after a historical individual whose death by drinking hemlock is described in Plato's Apology. In appearance, Socrates was not much to look at. He was snub-nosed, cared little about his dress . . .' Suppose then that the lecture about Socrates went in this way. Who was the lecturer talking about? Clearly, he was talking about Socrates the philosopher not Socrates the fishmonger.
whom he didn't yet know even existed. Was he also referring to Socrates the philosopher? Donnellan would be committed to the view that he was referring to Plato's uncle and making false assertions about him. But this is either quite absurd or it is introducing a notion of reference that does not include, in any way, the speaker's intention to refer to a unique individual.

Let us look at the first part of this dilemma. In a slightly different connection--i.e., in connection with referring uses of definite descriptions--Donnellan takes speaker's intention to be essentially involved in the act of reference; in fact, he takes it to be decisive in the determination of the referent. He argues that

when a definite description is used referentially, not only is there in some sense a presupposition or implication that someone or something fits the description, . . . but there is a quite different presupposition: the speaker presupposes of some particular someone or something that he or it fits the description.17

The example that Donnellan employs to illustrate his point is a case where one says at a party 'the man drinking the martini is happy', intending to refer to a man who is actually drinking a ginger ale. He assumes that it is clear from the context who is the referent: it can't be anyone but the man drinking the ginger ale. In this example, certain features of which are analogous to our Socrates example, the speaker's intention is decisive. He intends to refer to a particular individual and does so refer even though some of his beliefs about that individual are mistaken. It would seem that the same could be said in the Socrates example: The teacher is referring to the individual to whom he intends to refer. In a sense, we might say that the act of
reference on the part of a speaker is foolproof; his beliefs about whom he is talking or to whom he is referring are the last court of appeal. He may be misunderstood by his audience if they take him to be referring to someone other than the person he 'has in mind', but this would always be a case of misunderstanding the speaker's intention.

To return to our Socrates example, if the speaker's intentions are as crucial as Donnellan supposes in the case of referring uses of descriptions, it appears more reasonable to say that the lecturer was referring to Socrates the philosopher. If reference involves intention in the way just described, it would be absurd to claim that the lecturer is referring to Plato's uncle and making false assertions about him, since he knows absolutely nothing about Plato's uncle.

The second part of the dilemma concludes that Donnellan is supposing a notion of reference that does not involve speaker's intention. This does not seem likely given what Donnellan says about referring. He takes reference to be a speech act performed by a speaker—"the speaker uses the description to refer to something, and [I] call this the 'referential use' of a description"—and he even distinguishes this sense of reference from what Russell called the denotation of an expression—"... referring is not the same as denoting ...." It appears clear from Donnellan's example that Donnellan is raising a question about the referring use of the name 'Socrates' and not about the denotation of 'Socrates'.

In certain respects, the distinction between 'referential use' and denotation of an expression is similar to Grice's distinction between utterer's meaning and word meaning, between what a speaker meant and
what the word means. As Grice explains it, utterer's meaning is a function of the speaker's intentions, and it is basic. Sentence meaning is a function of utterer's meaning; i.e., it is explained in terms of 'meaning for U' where U is the speaker. Briefly, the meaning of a particular utterance-type $X$ in a particular context of its usage boils down to 'U has in his repertoire the procedure of uttering a token of $X$ if U intends the audience to believe such and such.' In this way the meaning of particular utterances is a matter of certain conventional behavior on the part of the speaker. For a language community in general, Grice defines what he calls "timeless meaning" as the meaning of an utterance-type independent of any particular context. This is defined in terms of a certain procedure in the repertoires of members of the linguistic community. Finally, word-meaning is explained in terms of the contribution which the word makes to the timeless meaning of the utterance.

To take an example, Grice raises the question of how a name like 'Fido' comes to have a particular object in the world as its denotation. Grice's answer, as we might expect, is that the naming relationship is one the speaker "intends to set up." 'Fido' denotes Jones' dog because it has been used to refer to Jones' dog and its use in this respect is intentional; that is, speakers intend to set up the relationship between 'Fido' and Jones' dog and other speakers adopt this convention as part of their repertoire. According to Grice, this shows that "intensionality is embedded in the very foundations of the theory of language." Our point, a much more modest one to be sure, is that denotation depends on speaker's reference and hence on speaker's intention.
Of course, it may be that Donnellan would not accept the Gricean analysis; he does not say one way or the other. But many of his remarks about reference appear to coincide with Grice's picture. After all, Grice's picture is an analysis of the speech act and Donnellan certainly assumes that reference is a speech act. How then would Grice's analysis apply to the lecturer's use of 'Socrates'? Clearly, it would entail that the lecturer was referring to Socrates the philosopher for, regardless of what may be subsequently discovered about Socrates, the lecturer was intending to refer to a certain individual that he had identified in a certain way, and his intention is here decisive.

Moreover, as for the denotation of 'Socrates', we can say that denotation is not something that words have independently of their use by speakers. More precisely, the denotation of 'Socrates' as it is currently used within a certain language community is dependent on the connections which speakers of that language intend to set up. It is at this point that the PID enters the picture, for the information believed about Socrates is what we depend on to set up this connection. To assume then that the denotation of 'Socrates' could be an unknown entity (Plato's uncle before the discovery of the diary) is to give words the power to confound our use of them. This goes against Strawson's dictum—which Donnellan embraces—that "referring is not something an expression does; it is something that someone can use an expression to do." To conclude then, we see that even if we employ a distinction between speaker's referent and denotation, we still have not succeeded in giving a clear sense to the 'theoretical possibility' that Donnellan urges on us.
A foreseeable objection to the above discussion would undoubtedly urge that I have begged the question by presupposing an intensionalist account of both speaker's reference and denotation when this is precisely what Donnellan (and Kripke) are attempting to challenge. In response to this it should be noted that the above discussion is not intended as a refutation of Donnellan's (or Kripke's) extensionalist approach to reference. We have yet to consider their positive accounts. The above discussion is intended in two ways. First, it attempts to show that Donnellan has failed to construct a counterexample to the PID. He has not succeeded in constructing an example for which the PID cannot provide a plausible, intuitively acceptable explanation. The key point involved here was that Donnellan's example constitutes an illegitimate broadening of the language community and thereby separates the act of reference from the speaker's intention. It will be seen shortly that even on their extensionalist account of reference both Donnellan and Kripke employ the notion of 'intending to refer' in a quite essential way. Therefore, it is not clear that I am begging the question since I am merely attempting to draw out the implication of an assumption that both Kripke and Donnellan make (and must make).

This is the second sense in which the above discussion was intended; namely, to spell out the implications involved in the view that reference is a speech act. It may be, but it appears unlikely that it is, that Donnellan's and Kripke's theories are not intended to explicate the notion of reference as a speech act. If not then they are not alternatives to the PID which does take this as its task. If so, then their theories are confronted with a serious problem. Can they explain how it is
possible for a language community to use a name to refer to an individual about whom nothing is known or believed? The above discussion was intended to show that this is unlikely.

Some Purported Counterexamples

Before we consider their positive accounts of reference, it should be noted that Kripke presents several purported counterexamples to the PID, all of which fail for reasons not dissimilar from those outlined above. We have already seen this in the case of his Feynman example, but it may be useful to look at one more of Kripke's examples.

In general, the problem with Kripke's examples is that they do not employ a complete or accurate picture of the PID. They tend to focus on certain features of the theory and argue that because such features lead to implausible results, the theory must be mistaken. However, when the features that Kripke calls attention to are viewed as part of an integrated picture, the apparent problems vanish. In short, Kripke's arguments are directed against 'straw men'; they have a certain attractiveness and prima facie plausibility which totally disappears when considered against a more complete picture of the PID.

The truth of these claims shows up no better than in Kripke's Gödel example. Kripke asks us to imagine "the following blatantly fictional situation":

Suppose that Gödel was not in fact the author of [Gödel's incompleteness] theorem. A man named 'Schmidt', whose body was found in Vienna under mysterious circumstances many years ago, actually did the work in question. His friend Gödel somehow got hold of the manuscript and it was thereafter attributed to Gödel. On the [PID], then, when our ordinary man uses the name 'Gödel', he really means to refer to Schmidt, because Schmidt is the unique person satisfying the description, 'the man who discovered the incompleteness of arithmetic' [that
is associated with the name Gödel]. So since the man who
discovered the incompleteness of arithmetic is in fact
Schmidt, we, when we talk about 'Gödel', are in fact always
referring to Schmidt. But it seems to me we are not. We
simply are not.25

Of course, we are not talking about Schmidt here but the relevant question
is this: Can the story be changed in such a way that we can be said to
refer to Gödel even though all the information believed true of Gödel--
the entire descriptive backing--is false? This is the question that
Kripke must tackle if he wants to construct a relevant counterexample.
As it is, his example draws on our intuitions about the described
situation—that 'Gödel' is used to refer to Gödel even though one par-
ticular description is false of Gödel—but does not attempt to account
for these intuitions. Clearly, however, these intuitions are explanable
by the PID, but only if we drop the unreasonable assumption that one
description constitutes the entire descriptive backing for 'Gödel'.

Once we allow the full picture of the PID to hold sway we see that it
does not imply that uses of 'Gödel' would, in that situation, refer to
Schmidt. They cannot because no speaker could intend to refer to an
individual about whom nothing is known. We also see that the PID would
imply that, in the above example, 'Gödel' would be used to refer to
Gödel; alternative means of distinguishing the referent could be relied
upon if one of the associated descriptions turns out to be false—e.g.,
'the former professor of mathematical logic' or, when push comes to
shove, 'the person who was born on such and such a date in such and such
a place'. This type of information is part of the background in which
the name 'Gödel' is used and it needn't be the case that everyone who
uses the name 'Gödel' be aware of these facts about Gödel so long as
they are generally available.
A second unreasonable assumption that is operative in Kripke's example is that reference is purely and simply a matter of satisfaction of certain descriptions. Actually, however, the picture suggested by the PID is slightly more complex because the speaker's reference is always a matter of belief and intention. Therefore, strictly speaking, we do not refer to objects that satisfy (simpliciter) a certain descriptive backing. We refer to objects that are believed to satisfy certain descriptions. In most cases, quite by accident, the distinction is not necessary, but sometimes it makes a crucial difference, as in Donnellan's example about the man with the martini. In that case the speaker was referring to a man whom he believed to have a martini. The fact that this belief was mistaken turned out to have no effect either on the success of his reference or on his reference being understood. It would have been misunderstood if the hearer had believed that some other individual present who had a martini was the intended referent. However, in most cases of referring uses of names, the difference between 'satisfied' and 'believed to satisfy' makes no (operational) difference.

One reason that this difference is not always relevant is that there is often no effective procedure for deciding when the intended referent fails to satisfy the associated descriptions; most of the purported cases where such failures occur are thought-experiments where we're asked to assume, for example, that Socrates was not Plato's mentor. In such cases reality is given to us as something independent of our beliefs: we believe Socrates to be such and such when in reality he is not—or so the argument goes. But clearly, the kind of 'reality' that forms the background of actual references to individuals like Socrates is a function of our beliefs, not something independent of them.
This point can perhaps be sharpened somewhat by employing Russell's

distinction between knowledge by acquaintance and knowledge by description

(in a fashion different from that in which Russell employed it).\textsuperscript{26} If we
drop Russell's assumption that acquaintance is always acquaintance with

sense data, then we might see acquaintance as a continuum, a matter of
degree. I am more acquainted with an object that is in my field of

perception than one which is not; I am more acquainted with an object

that is a possible object of perception than one which is not; and so

forth. The continuum can be viewed as shading into those objects which
can be known only by description; i.e., those objects which we have not

perceived and cannot possibly perceive. Clearly, we can arrange the

intended referents of particular uses of denoting expressions along such

a continuum; in terms of the expressions involved the continuum would
range from expressions like 'that dog there (pointing)' to those like

'the center of mass of the universe'. Also, it is clear that the identity

conditions may vary along the continuum in the following sense: If I
intend to refer to that dog there with the name 'Fido', there is the
possibility that the gap between my belief that that dog is Fido and the
reality that in fact that dog is not Fido can be settled effectively; by

'effectively' here I mean it can be settled by reacquaintance with the
proper bearer of the name 'Fido'. For example, suppose Fido is Jones'
dog and Jones comes along and says 'that dog is not Fido; this dog
(pointing) is Fido'. In such a case the problem of mistaken identity
is settled right on the spot. Moreover, in such cases it makes sense to
say that I was referring to the dog that I believed to be Fido but which
wasn't in fact Fido. The facts at this level can be settled effectively
and are quite independent of my beliefs and my audience's beliefs.
Things are a bit different—but only by degree—at the other end of the continuum where the facts and the objects are known only by description. For example, when I use the expression 'the center of mass of the solar system' in a referential way—e.g., 'The center of mass of the solar system is situated at the center of mass of the sun'—there is no possibility of correcting a mistaken belief on my part by reacquaintance with the real center of mass of the solar system. Mistaken beliefs, in this case, must be corrected by re-describing the center of mass of the solar system. (As a simplifying assumption, assume that I was not referring to some coordinate on a topological model of the universe, in which case it would make sense to correct the mistaken belief by reacquaintance with a different coordinate.) The case of such abstract definite descriptions is relatively transparent. Proper names for individuals that fall towards the abstract end of the spectrum are similar but not quite so transparent. For example, would we say that someone using the numeral '5' referred to something he believed to be the number 5 but which wasn't in fact the number 5? It seems unlikely. What about a name like 'Socrates' where the identity conditions are not quite as fixed as those of the number 5? Socrates is an individual that we know only by description; some of his contemporaries were more directly acquainted with him. But the question is how do we refer to Socrates? There is, or was a real Socrates but we are not acquainted with him. We know him only by description which is to say that as far as our use of language is concerned, Socrates is as we believe him to be. Or, more precisely, he may have been different than we believe him to be, but the only way we could state this is by employing other
descriptions than those we currently hold true of Socrates. In effect then the question of settling the identity of Socrates is a little like rebuilding Neurath's ship while it is still afloat: We remove a plank here and replace it with another, and perhaps, we eventually replace every plank in this way; but the question is moot as to at which point we have described the real Socrates. In effect then, the distinction between the real Socrates and the person we believe to be Socrates collapses; it does not affect our ability to refer to Socrates.

To tie this back into Kripke's Gödel example we might say the following: Even if we assume that Gödel is known or described only in terms of the description 'the discoverer of Gödel's theorem', so that the entire descriptive backing of the name Gödel is false, this still would not constitute a counterexample to the PID. For, in that case, we could have been referring to the man believed (by everyone in the language community) to be the discoverer of Gödel's theorem, and that man is Gödel, not Schmidt. Insofar as the knowledge that Gödel is an imposter is not part of the corpus of beliefs presupposed in the context of the speech act, it has no effective role to play in determining the referent.
1. See above pp. 97 ff.


3. Ibid.

4. Ibid.

5. Ibid.


8. Ibid., p. 229.


10. Ibid.

11. This, at least, is how Searle, Dummett, Strawson and other proponents of the PID conceive the theory.


13. Ibid.


15. Ibid., pp. 138-139.

16. This was pointed out to me by Dr. Copi.

18. Ibid., p. 43.

19. Ibid.


21. Ibid., p. 233. Actually, Grice's treatment is much more complex and much more general than described here. It applies not only to informative discourse (the lecture on Socrates) but to commands, questions, etc. To make my point I have simplified Grice's analysis but this should not be taken as anything more than a bare sketch of Grice's theory.

22. Ibid., p. 239.

23. Ibid.


An Exposition of Donnellan's and Kripke's Theories

Let us turn now to Kripke's and Donnellan's positive account of reference. Their individual approaches to proper names, which were developed independently, have much in common and can be considered together. We shall point out what slight differences exist in the course of the discussion but these will not affect our argument in any way. We begin with a brief exposition of each account.

Donnellan calls his theory 'The Historical Explanation Theory' of reference (HET), and sets out its "main idea" as follows:

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.¹

Donnellan admits that his thesis could stand clarification and that, in general, the HET has not been developed to the point where it can provide "a completely detailed answer" of the appropriate relationship between an act of using a name and its referent.² One thing does emerge quite clearly from this statement of the thesis and that is that the HET is a thesis about what we have called speaker's referent. The referent of a particular use of a name is the object that enters into a historically correct explanation of who or what is the object intended by the speaker. By way of clarification, it should be noted that "successful reference"
does not imply that the speaker's reference be understood by the audience; it implies only that the speaker has successfully referred to some unique individual or other.

The key notion in Donnellan's theory is that of a "historical explanation" which he explains thus:

we search not for an individual who might best fit the speaker's descriptions of the individual to whom he takes himself to be referring (though his descriptions are usually important data), but rather for an individual historically related to his use of the name 'Socrates' on this occasion. 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 modelled 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'.

According to Donnellan such an explanation could yield an individual as the referent of 'Socrates' "even though that individual is not correctly described by the speaker's attempt at identification." One problem with putting things this way, as we mentioned above, is that it places too much weight on the speaker's ability to identify the referent and not enough weight on the social character of the use of names. Like all other words, proper names have an established usage which does not depend on the competence of any particular speaker; but this social character of language-use has not been passed over by proponents of the PID. In large part, the notion of a 'descriptive backing' is intended to relieve the particular speaker from the burden of uniquely identifying the referent of 'Socrates'. Enough has been said about this both here and by others and we needn't go into it again. However, to forestall confusion about what is at issue between the PID and Donnellan's HET we can
make the following qualification of the HET: For the HET to count as a
genuine alternative to the PID, it must establish not only that the
referent picked out by the historical explanation is different from the
individual described by some particular speaker's attempt at identifica-
tion, but also that the referent is different from the individual picked
out by the descriptive backing that is associated with the name and
operative within a certain community of speakers. This reformulation of
Donnellan's claim should be uncontroversial since it does appear to
spell out more clearly what is an important implication of his theory.

Donnellan appends a qualification to his view which is concerned
with the key notion of the 'omniscient observer': He admits that
ordinary speakers neither know "the detailed history behind the uses of
names" nor do they "often make the [sort of historical inquiries which
would] reveal those details."5 He admits that we usually just assume
that if another speaker's descriptions of the referent square pretty
much with our own, then the other speaker is referring to the same
object as we are. According to Donnellan, all the HET need maintain is
that "the final test for reference is the kind of historical connection
I have described, [and] that the customary assumptions and use of
indicators are in the end dependent upon being fairly reliable guides
to the existence of such a connection."6

To prove the cogency of the theory, Donnellan sets himself the
following conditions:

[(i)] when there is an absence of historical connection
between an individual and the use of a name by a speaker,
then, be the speaker's descriptions ever so correct about
a certain individual, that individual is not the referent
...
[(ii)] a certain historical connection between the use of a name and an individual can make the individual the referent even though the speaker's descriptions would not by themselves single out the individual.7

Donnellan also points out that the type of historical explanation alluded to is not the history of the use of the name--because Socrates may not have been called by the name 'Socrates'--and that 'historical' is used in 'the broadest sense possible.'8 We shall consider some examples and some further clarification of Donnellan's theory when we come to criticize it, but for now the above should suffice as a reasonably complete picture of the HET.

Kripke calls his theory the 'Causal Theory' (CT) and the key difference between his and Donnellan's theory is that Kripke views the link between the use of a name and its referent as a causal chain rather than a historical explanation. However, this difference may be largely terminological, since Donnellan compares his theory to the causal theory of perception.9 We shall not pursue this question here.

Kripke balks at calling his theory a 'theory of reference' and prefers that it be viewed as a picture of how reference takes place. His "rough statement" of this picture goes as follows:

An initial baptism takes place. Here the object may be named by ostension, or the reference of the name may be fixed by a description. 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.10

One difference between Kripke's and Donnellan's theories concerns the notion of an initial baptism. Donnellan drops this requirement because "it is not a theoretical necessity that names enter our linguistic transactions in this way."11 Donnellan is undoubtedly right here, at
least if we take 'baptism' in its ceremonial sense, for clearly, names can become attached to individuals without any such formal procedure. Therefore, it is best, perhaps, to drop this feature of Kripke's picture. It suffices for either theory that the initial link in the chain—the link between the name and the individual—be effected in some way or other.

A somewhat fuller statement of Kripke's view is given in 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 one at the far end of this chain, who has heard about, say Richard Feynman, in the market place or elsewhere, may be referring to Richard Feynman even though he can't remember from whom he first heard of Feynman or from whom he ever heard of Feynman. He knows that Feynman was a famous physicist. A certain passage of communication reaching ultimately to the man himself does reach the speaker. He then is referring to Feynman even though he can't identify him uniquely... So he doesn't have to know [how to identify Feynman uniquely], 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...12

We see then that the key concept of both Kripke's and Donnellan's theories is that of some kind of link between the use of a name and its bearer and not a set of identifying descriptions.

**Genuine Alternatives to the Intensionalist Approach?**

The first question that must be raised concerning these theories, which we shall abbreviate as HET-CT when concerned about their common features, is in what sense and to what degree do they differ from the PID. Are the PID and the HET-CT genuine alternatives? It has been maintained throughout this discussion that the PID should not be construed (as both Kripke and Donnellan tend, at times, to construe it) as a theory which
is blind to the social aspect of successful reference. Several of its proponents have stated the PID in terms of what must be presupposed by some particular speaker in order for successful reference to occur, but this is the result more of the methodological assumption that knowledge about reference in general can be gained by looking at a particular speech context involving a speaker and a hearer. It does not follow from this methodological assumption, and it is not part of the theory itself, that in order to employ a name successfully each speaker must be able to state sufficient identity conditions for its referent. In one way or another all of the proponents of the PID have acknowledged the phenomenon of a parasitic employment of names, dependent on the knowledge or belief that someone in the language community can provide identity conditions. This 'parasitic reference' follows from the fact that names, like all other words, have a common use in a language.

Yet, in spite of this, it appears that the main selling point of both the HET and the CT is their cognizance of and emphasis on the social aspect of language. Can it be—in this post-Wittgensteinian era—that the proponents of the HET-CT believe they've discovered something that was overlooked by the traditionalists? No, for even Kripke acknowledges that he shares his insight, in part, with Strawson. Yet in most every example of and purported counterexample to the PID that they employ, Kripke and Donnellan represent the PID as pertaining to descriptions presupposed by some particular speaker; indeed, they appear to take this as the major fault of the PID and the reason for devising their alternative theories.

However, as is perhaps evident already, neither Donnellan nor Kripke is prepared to abandon completely the notion of a descriptive backing.
Donnellan considers it as a "guide" to the existence of the historical chain whereas Kripke sees its "primary applicability" as "that of the initial baptism." Nevertheless, I suspect that the role played by some sort of descriptive backing is far more crucial to their own approach than either Donnellan or Kripke has indicated. We shall pursue this point in due course but for now we shall ask: Is it the case that the HET-CT and the PID differ merely in emphasis since both appear to acknowledge the social aspect of the use of names and both appear to involve appeal to a descriptive backing for names?

Kripke has raised this same question in regard to Strawson's remark that "one reference may borrow its credentials, as a genuinely identifying reference from another; and that from another [without fear of an infinite regress]." Kripke acknowledges that indeed Strawson's insight is a good one and quite similar to the raison d'être of his own theory. He claims however that Strawson's theory reveals "a difference at least in emphasis from the picture I advocate, since he confines his remark to a footnote. The main text advocates the cluster-of-descriptions theory." This is an unusual sort of claim, for it suggests that Strawson's insight is not a part of the description theory when in fact it is an important part. As for the comparative lack of emphasis on Strawson's part, this may be due to several factors.

First, it may be due to the fact that the phenomenon referred to in Strawson's note—and in Kripke's theory—is so commonplace as to require little discussion. Surely, Strawson—and the other proponents of the PID—is not blind to the fact that names are passed from speaker to speaker. Secondly, the main concern of Strawson, and the PID in general,
is to explain how it is that we are able to understand references to individuals. Towards this end Strawson simplifies the picture by analyzing the act of reference in terms of a situation involving just one speaker and one hearer. Because of this, the theory is stated in terms of what the individual speaker must presuppose in order to refer successfully. But, clearly, this is just a feature of Strawson's method, not his theory. His theory is meant to be a general account of reference applicable to a language community as a whole. Also, because, methodologically at least, he views reference from the perspective of the individual speaker and requires that the speaker must be able to produce an identifying description, Strawson must give some account of why it is that all speakers cannot do this for some names. His answer is that we rely on another's ability to identify the referent. But this question of 'parasitic language use' is less important theoretically than that upon which it is parasitic. The more important question, and that to which the PID brings genuine explanatory power, is how one is able to refer to a unique individual and be understood as referring to that individual, assuming the reference is not parasitic on that of another speaker. In view of these considerations, it is unlikely that much can rest on Kripke's claim of a difference in emphasis; if that is the only difference, it's not clear we have alternative theories here.

Kripke claims that there are substantive differences between his and Strawson's theories as well:

Just because Strawson makes his remark in the context of a description theory, his view therefore differs from mine in one important respect. Strawson apparently must require that the speaker must know from whom he got his reference, so that he can say 'By "Gödel" I mean the man Jones calls "Gödel"'.

If he does not remember how he picked up the reference he cannot give such a description. The present theory sets no such requirement. As I said, I may well not remember from whom I heard of Gödel, and I may think I remember from which people I heard the name, but wrongly.¹⁶

Not only is Kripke misled by appearances here, but he has involved himself in what seems, at first blush at least, an inconsistency. He has stated of his own theory that when one "receives" a name from another speaker, one "must, I think, intend when he learns it to use it with the same reference as the man from whom he heard it."¹⁷ But how can one employ a name with the same reference as another without remembering from whom he acquired the name? Perhaps Kripke can avoid this problem by arguing that the receiver of the name need only remember how it was employed and not from whom it was acquired. However, it is not clear that remembering how a name was employed does not involve remembering the backing of descriptions with which it was employed. Unless Kripke can show that it is possible to remember how a name is used without appealing to some backing of descriptions, he has failed to disassociate his theory from the PID. In order to "intend . . . to use it with the same reference" one must be able to state some sort of identity conditions. That is, one must be able to produce on demand some description or other of the referent, unless of course, Kripke means "same reference" in some trivial sense like, say, 'By "Cicero" I intend the same reference that Jones intended whatever that may be'. Not only does this explanation of 'same reference' involve one in a regress that must end with some sort of description of the referent itself, but it involves remembering from whom one acquired the name.
Secondly, even in cases of parasitic reference, the description theorist is not committed to descriptions of the sort Kripke supposes: viz., "By 'Gödel' I mean the man Jones calls 'Gödel'". The description theorist would be committed to reliance on such descriptions only if it were the case that names could be passed from link to link without also passing along some information about the referent. This does not appear to be the case however, although in several of his examples Kripke appears to think so. Consider a typical example:

Suppose that the speaker has heard the name 'Cicero' from Smith and others, who use the name to refer to a famous Roman orator. He later thinks, however, that he picked up the name from Jones who (unknown to the speaker) uses 'Cicero' as the name of a notorious German spy and has never heard of any orators of the ancient world.\(^\text{18}\)

Kripke concludes that his theory would succeed in picking out Cicero as the referent because his theory is interested in "the actual chain of communication," not the misremembered chain as Strawson's theory must suppose (according to Kripke).\(^\text{19}\)

Like so many of Kripke's examples this one suffers (but gains whatever plausibility it might have) from an underdescription of the speech context. Is it really plausible that the speaker in question could employ the name 'Cicero' in some future context without knowing a single thing about Cicero except where he might have heard the name? And what then could he say about Cicero? Maybe the speaker heard Smith just simply saying 'Cicero' and nothing else about Cicero? I seriously doubt that in such cases we would want to say that the speaker can use 'Cicero' in a significant way. As an intuitive test that may shed some light on the matter, we might try to think of a proper name for someone or
something about which we know absolutely nothing and then try to say something significant about that thing. To my mind we'd be uttering nonsense.

A far more reasonable picture of how the name got passed on would involve supposing that the speaker heard at least one fact about Cicero and can rely on this to explain to whom he is referring when he uses the name. Moreover, the description theorist is not committed to the view that he must rely on a description like the one above because, again, he can exploit the fact, known to the speaker, that the name is part of a common language; the speaker can assume that someone can uniquely describe the referent.20

As the above argument shows, names do not get passed along in a vacuum and therefore, it is incumbent on speakers, if they are to employ a name correctly, that they obtain some information about its bearer. It follows that Kripke cannot confine the applicability of the backing of descriptions to the "initial baptism". To see this, we can consider how it is that a theory like Donnellan's or Kripke's would distinguish between a speaker's use of 'Cicero' as a name for a Roman orator and its use as a name for the German spy. This would present no problem for the PID which can rely on differences in the descriptive backing and differences in the speaker's intention to distinguish the two uses. What sort of explanation can the HET-CT provide?

Assume then, the existence of two distinct chains leading from the use of 'Cicero'; one goes to the orator and the other to the German spy. To remove contextual indications assume that the context in which the name is used is discontinuous in space and time from any previous
use of the name and further that the use in question starts out as follows: 'Cicero was really a very moral person. He took up the profession that he did because he was deeply committed to his country and to his fellow citizens; a true patriot was Cicero.' Now, assuming the existence of the appropriate chains—or historical explanations—to which of the possible Cicero's is the speaker referring? Clearly, only the speaker's intention to refer to a unique individual can decide the question. By 'intention' here we do not mean that the use of 'Cicero' is accompanied by a mental picture of Cicero. We mean only that when asked to whom he was referring the speaker would respond somewhat as follows: 'Why to the famous Roman orator, of course.' We can see from this example, that it is not so much the existence of the appropriate chains that determines the referent, but the fact that the speaker can provide identity conditions for the object to whom he refers. I do not mean to suggest that either Donnellan or Kripke is barred from handling the problem in this way. Donnellan explicitly acknowledges the role that descriptions play as "guides" to the appropriate chains, and it appears that Kripke must also acknowledge something of this sort. The point is this: if the presence of the sort of descriptive backing relied on here is the crucial factor in this situation, why isn't it the crucial factor in all (referring) situations? Surely, the role played by the descriptive backing is not merely disambiguation of ambiguous names; it appears to be the genuine explanatory answer to the question 'To whom is the speaker referring by "Cicero"?'
Some Problems with Donnellan's and Kripke's Proposals

It appears, then, that something like the descriptive backing of the PID must be incorporated into the HET-CT. The question is to what extent? Is it merely supplemental, as Donnellan suggests, or does it have a more crucial role to play? These are not easy questions to answer without begging the question at issue. For, it would appear that any answer would presuppose one theory or the other. But perhaps the adequacy conditions set out by Donnellan can provide us an entering wedge. That is, (i) can it be shown that a name can be used to refer to an object in the absence of the historical (or causal) chain of the sort required by the HET-CT? And (ii), can a name be used to refer to an individual linked by the appropriate chain to the name even where this link is generally unknown in the language community? If the description theorist can show that (i) may obtain but that (ii) may not, this would constitute evidence in favor of the PID. Conversely, if the HET-CT theorist can show that (ii) may obtain while (i) may not, this would constitute evidence in favor of Kripke's and Donnellan's approach. Let us then conduct the following thought experiments.

First, can a name be used to refer to an object in the case where a historical connection does not exist between the name and the object? Gareth Evans has proposed several examples which address this question, all of which show that an affirmative answer is possible. He concludes that "[t]he causal origin of the speaker's familiarity with the name, save in certain specialized 'mouthpiece cases' [e.g., relaying a telephone message], does not seem to have a critical role to play." Perhaps the most convincing of Evans' examples attempts to show that the
chain can lead back too far. For example, assume that somewhere along the line an imposter had killed Napoleon and adopted his name. Then the chain would lead back to two individuals, somewhat as in the following diagram:22

```
<table>
<thead>
<tr>
<th>imposter</th>
</tr>
</thead>
<tbody>
<tr>
<td>'Napoleon'</td>
</tr>
</tbody>
</table>
```

In such cases there is a historical connection—so this example does not directly address the above condition—but the chain does not pick out a unique individual. (In this respect, the example is similar to the Cicero example considered above.) How then are we to decide to whom the use of the name refers?

Evans' thesis is that we must rely on information about the referent that is cullable from either the context or the speaker. However, he argues that information, per se, is not enough, for in this case the information would not distinguish the real Napoleon from his imposter. Therefore, Evans argues, we must also consider the source of the information. For example, assume that the imposter entered the scene in 1814—i.e., after Napoleon had done most of the things for which he is duly famous. In that case, Evans contends, most of the historical information associated with 'Napoleon' would be dominantly of the real Napoleon. On the other hand, if the imposter took over at a much earlier point in Napoleon's history, say 1793, then the associated information would be dominantly of the imposter. He would have been responsible for the feats for which Napoleon is famous. Consider, then, the sentence 'Napoleon was defeated at Waterloo in 1815.' According to Evans, if the imposter took over in 1814, historians would be said to
have false beliefs "about who was the general at Waterloo."\(^{23}\) But if the imposter entered the scene in 1793, historians have correct beliefs about who was general at Waterloo—it was the man who was famous for certain exploits between 1793 and 1814—but false beliefs about that man's earlier career.\(^{24}\) Thus, according to Evans' compromise solution, we must rely on both information about the referent and the source of that information in order to determine the referent of a proper name.

According to Evans, the key point of his thesis is that the fix between a use of a name and its referent is due ultimately to its causal origin and not by whether the name fits certain associated descriptions.\(^{25}\) One problem here, however, is that the description theory would yield precisely the same conclusions as Evans' compromise. To see this consider the Napoleon example again. If the imposter took over in 1793, he would be the dominant source of (most of) the relevant information and hence the referent of 'Napoleon' in 'Napoleon was defeated at Waterloo'. If he took over in 1814, he would not be the dominant source and hence not the referent of 'Napoleon' in that sentence. According to the PID, if the imposter took over in 1793, most of the information associated with 'Napoleon' would be true of the imposter and therefore he would be the referent of 'Napoleon' in the above sentence; if he entered in 1814, the associated information is false of him but true of the real Napoleon; hence the name would refer to the real Napoleon. Thus, the Napoleon example does not support the conclusion that the "fix" is determined by source rather than fit. Evans' conclusion—that, in general, a speaker intends to refer to the item that is the dominant source of his associated body of information\(^{26}\)—is a bit too
hasty. Moreover, it looks very much like the sense in which we can say that an object is the source of information is more a matter of what information is true of that object—which is a question, really, of fit or satisfaction—than any kind of causal relationship. Evans has failed, really, to distinguish his theory from the traditional PID.

Can it be shown that source—in a causal sense—rather than fit is the crucial factor in determining the referent? If the source were crucial, it should be impossible for a name to be used to refer to some object that is not historically (or causally) connected to the name. So we are back to condition (i) above. Consider the following example, which is one of Evans' though I'm not sure he would draw the same conclusion from it as I wish to draw.

An urn is discovered in the Dead Sea containing documents on which are found fascinating mathematical proofs. Inscribed at the bottom is the name 'Ibn Kahn' which is quite naturally taken to be the name of the constructor of the proofs. Consequently it passes into common usage amongst mathematicians concerned with that branch of mathematics. 'Khan conjectured here that . . .' and the like. However suppose the name was the name of the scribe who had transcribed the proofs much later; a small id scripsit had been obliterated.27

Evans concludes that 'Ibn Kahn' would be one of the names of the mathematician because "there is a coherent community using the name with the mathematician as the intended referent."28 But this certainly does not uphold his contention that the "fix" between a name and its referent depends on the source of the information rather than the fit. The source of the information in this case, if it can be called the source at all, appears to be the assumptions of the mathematicians that the name is the name of the author of the proofs. The information which they associate with the name—e.g., 'Ibn Kahn authored such and such a proof; he must
have been quite a genius to do such advanced work, etc. — is perhaps true of the author but it certainly does not originate with the author in any significant sense. It would seem, therefore, that again, the crucial factor in determining the referent is both the associated information and the appropriate intentions. Moreover, we have here an example in which a name is used to refer to an individual (the mathematician) who is neither historically nor causally connected with the use of the name. (Except, perhaps, in a rather ad hoc sense which we shall consider presently.) An omniscient observer would see that the actual historical chain leading from 'Ibn Kahn' would lead to the scribe. So on the HET-CT view, the mathematicians are referring to the scribe and making false statements about him—a most counterintuitive conclusion.

It might be argued that the above example does not constitute a counterexample to the HET-CT because it can be claimed that the mathematicians' speech behavior amounts to a dubbing of the author with the name 'Ibn Kahn' and therefore constitutes the first link in the chain. The problem with this sort of reply is its very ad hoc character. Clearly, the same type of response could be given to every problem that we throw up in the face of the HET-CT. That is, if this sort of reply is acceptable, there could be no possible case in which a name referred to an object with which it wasn't linked; we can always supply the link-up in this ad hoc fashion. If such an ad hoc response does not offend the letter of the HET-CT, it certainly goes against the spirit of the theory. For one thing, the omniscient observer is no longer needed to settle problems of reference; our ad hoc explanation could overrule him in every case just as in the present example. Secondly, the ad hoc explanation
comes dangerously close to acknowledging that certain names are just disguised descriptions. For in this case the name 'Ibn Kahn' is very closely linked to a certain set of descriptions believed by the mathematicians to be true of the author of the proof. Therefore, although the HET-CT theorist can avoid the conclusion we have drawn here, he does so at some cost. I should think that until a less ad hoc explanation can be provided for the above example, the HET-CT theorist has failed to show that the first condition, (i), above, cannot obtain.

What about condition (ii)? Can a name be used to refer to an object historically connected to a name but where the connection is unknown to the language community? We have already dealt with this question in the preceding chapter and we concluded there that it appears very unlikely that (ii) may obtain. However, we agreed to postpone judgement so as not to beg any questions. Now that we have gained some insight into Donnellan's and Kripke's program we can construct an example in which (ii) obtains, explains it in terms of their theory, and if it leads to no absurdities we can admit that (ii) may obtain. Thus, consider Donnellan's Socrates example again. Suppose the individual whom an omniscient observer would pick out as the referent of 'Socrates' in 'Socrates was snub-nosed' is Plato's uncle, the fishmonger. This, again, is a case where the use of a name refers to an individual to whom it was not intended to refer. So every time Greek scholars or philosophers intend to refer to Socrates the philosopher they are really referring to Plato's uncle, and most of what they say with the name 'Socrates' is, strictly speaking, false. To my mind, this consequence is patently absurd. It turns reference into something magical over which speakers
have absolutely no control. For example, it would appear to render obsolete such questions as 'To whom are you referring?' and such answers as 'I don't care what they say, I was referring to Socrates the philosopher when I used the name "Socrates"'. It would no longer matter what speakers say about their intentions since it is historical chains that determine reference.

I'm not sure how Donnellan and Kripke would respond to these objections. Nor do I wish to suggest that they are blind to the importance of speakers' intentions in the act of referring. But the above problems do appear to follow from their theories. Perhaps the HET-CT approach can be reformulated in some way that avoids these conclusions; but it seems to me that any such reformulation would involve appealing to associated information, intentions and in short, to precisely those notions which descriptive theorists have claimed are so crucial. Whether the HET-CT theory can be amended in such a way as to avoid intensionalism seems unlikely, but it goes without saying that the burden of proof rests on the proponents of that view.
NOTES--CHAPTER VIII


2. Ibid.

3. Ibid., pp. 229-230.

4. Ibid., p. 230.

5. Ibid., p. 231.

6. Ibid.

7. Ibid.


13. Ibid., p. 349, n. 42.


15. Kripke, p. 299.

16. Ibid.

17. Ibid., p. 302.

18. Ibid., p. 299.

19. Ibid.

20. See above, p. 163.


22. Ibid., pp. 207-208.

23. Ibid.
24. Ibid.
25. Ibid., p. 214.
26. Ibid.
28. Ibid.
CHAPTER IX
CONCLUDING REMARKS

In the above discussion we have taken a detailed look at recent proposals in the theory of reference. Briefly stated, the upshot of this discussion has been that sufficient reasons for abandoning the principle that intension determines extension have not been adduced by the proponents of these recent proposals. Once we disassociate the notion of an intension from the conceptualistic and essentialistic interpretations, it provides a suitable basis for an account of the referential component of language. For the most part, the focus of the discussion has been the negative one that certain theories and approaches are problematic and, perhaps, mistaken. In closing, then, it may be appropriate to dwell, briefly, on the more positive implications of our conclusion—viz., on what is right about the intensionalist approach.

First, as we have seen, the recent proposals by Kripke, Putnam and Donnellan have all been directed in one way or another at the principle that intension determines extension. What we have seen is that their attacks on this principle—both as regards its application to proper names and to general terms—have consistently been found wanting. The consistent failure of their attacks to detract from the plausibility of the intensionalist principle reaffirms what we said at the outset: viz., the notions of intension and extension are inextricably bound together; they are correlative notions. This is the fundamental insight upon which intensionalism must be based, and, as we have seen, it is
this insight that emerges intact once we disassociate intensions from their conceptualistic and essentialistic interpretations.

Secondly, as we noted at the outset, the intensionalist approach provides a basis for characterizing the abstractive, inductive elements that make up the referential component of human language. As regards general terms like 'human', our ability to apply the word correctly involves not only the knowledge that it has been applied to certain objects. It also involves the knowledge that the term would be correctly applied to objects similar to those to which it has been applied. It would be mistaken to regard one as 'knowing the meaning of "human"' if he did not possess this knowledge. And we need not belabor the point that this knowledge involves the ability to abstract those features by which humans are similar. There are many such features and it matters little which particular features are invoked so long as they pick out the right objects. Similarly, with regard to proper names and singular reference, the ability to refer to an individual involves the ability to identify that individual and distinguish it from others. Again, we are led to reliance on features or characteristics of the individual and, if we are to be understood by our fellow speakers, we are led to reliance on features that they are also capable of knowing.

In either case, we might view the reliance on such an abstraction of features as, really, the very basis of the use of symbolism. It is important to note how inadequate extensions alone are toward this end. Even if we could display a multitude of the objects in the extension of 'human', we have still failed to convey the meaning of 'human' if we have not conveyed that 'human' refers to the objects in the display and
all those similar in certain respects. And we cannot convey this unless we mention attributes or features or characteristics of the objects. Thus, the mistake in the recent proposals by Kripke, Donnellan and Putnam is not so much their emphasis on extensions and on actual chains of reference, as it is their lack of appreciation for the importance of intensions as the basis for an understanding of the abstract nature of symbolism. In short, because the connections between words and things are human contrivances, and because these connections are maintained through constant human interplay, there is no getting away from reliance on intensions if we are to understand the connection between language and the world; for intensions, as we have seen, are the carriers, in abstract form, of human knowledge and invention.

In light of the above considerations then, my conclusion is that we must rely on both intensions and extensions in order to explain the referential aspect of language: extensions provide us with an answer to the question 'what do words refer to?' and intensions, by describing or characterizing the extension, provide us with a means of conveying this answer. To conclude, it is a recognition of the close connection between intension and extension, however imperfect that recognition may have been at times, that constitutes the basis of the traditional claim that intension determines extension.


Carnap, Rudolf. 'Meaning and Synonymy in Natural Languages,' *Philosophical Studies*, 7 (1955), 33-47.


Davidson, Donald. 'Truth and Meaning,' *Synthese*, 17 (1967), 304-323.


Donnellan, Keith. 'Proper Names and Identifying Descriptions,' *Synthese*, 21 (1970), 335-358.


Kripke, Saul. 'Semantical Analysis of Modal Logic I,' *Zeitschrift fur Mathematische Logic und Grundlagen der Mathematik*, 9 (1963), 67-96.

Kripke, Saul. 'Semantic Considerations on Modal Logic,' *Acta Philosophica Fennica*, 16 (1963), 83-94.


Pak, Ty. 'Semantics and Grammar: A Review of Recent Theories (Part II),' Semiotica, 13 (1975), 41-96.


