A GRICEAN THEORY OF REFERENCE

by

MICHAEL DOUGLAS BEEBE

B.Sc. Oregon State University
M.A. Simon Fraser University

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

in the Department
of
PHILOSOPHY

We accept this thesis as conforming to the
required standard
Jonathan Bennett
D. G. Brown
Thomas E. Patton

THE UNIVERSITY OF BRITISH COLUMBIA
August 1974
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the Head of my Department or by his representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Philosophy

The University of British Columbia
Vancouver 8, Canada

Date August 30, 1974
Abstract

I propose to analyse referring, the typical function of proper names and definite descriptions, in terms of speakers' intentions rather than in terms of the meanings of words. Grice's theory of meaning explains how a speaker can mean something simply by making an utterance with the proper sort of intention, and I attempt to apply this theory of meaning to referring. I see two related reasons for thinking a Gricean theory of reference correct. First, I try to show that our speech practice is such as to demand a Gricean-type theory; we simply do depend rather often on hearers' recognition of speakers' intentions to achieve reference. Second, the causal theory of proper names, which I investigate in Chapter 5, seems to demand a Gricean theory of reference.

In Chapter 1, I am concerned with whole utterance meaning, and with the problem of going from occasional meaning to conventional meaning. Lewis' analysis of convention is deployed in a way which I think vindicates Grice's hope that conventional or timeless meaning may be analyzed in terms of occasional meaning. In Chapter 2, I attempt to extend the Gricean program to parts of utterances. That is, I argue that an utterance may have meaningful parts and grammatical structure entirely without benefit of convention. Primitive cases of referring and predicating can arise at this pre-conventional level, and here certainly referring is to be explained
by a Gricean theory. Chapter 3 contains the body of my argument for a Gricean theory of reference. I try to show that our referring practice is in fact one which depends on the characteristic form of the Gricean intention: the speaker, by his utterance, intends that the hearer shall be affected in some way, intending this to come about at least in part by hearers' recognition of speakers' intentions. The first three section of Chapter 3 are devoted to arguing for the Gricean character of reference, and the remaining two to developing suitable epistemic foundations for such a theory. If a speaker's intention is primary, the meanings of the words he uses are (with qualifications) secondary, so we must be able to explain how a speaker is connected with the object of his reference in some way which does not essentially involve his being able to say something true about that object. I use Kaplan's theory of relational belief, with its emphasis on the causal element in a belief, to provide this connection. Having a relational belief about an object connects one with it, and allows one to have an intention directed towards it. Chapter 4 is a criticism of the standard theory of reference with Searle and Strawson cast as its defenders. Chapter 5 is a presentation and elaboration of the 'causal' or the 'historical explanation' theory of names, the new theory of names which, I claim, requires a Gricean theory of reference.
Acknowledgments

I should like especially to thank my thesis advisors Thomas Patton and Jonathan Bennett for their criticism and encouragement throughout the several drafts of this thesis. I owe thanks also to D. G. Brown and to Howard Jackson for their help with particular problems.

Also, I here acknowledge four years of support from the Canada Council, which gave me the time and financial security to write.
# Contents

## Introduction

1. Meaning and Convention  
   - Introduction  
   - A Nominalist Strategy for Part-Meaning and for Grammar  
   - Occasional Part-Meaning and Occasional Grammar  
   - Conventional Part-Meaning and Conventional Structure: Words and Grammar

2. Intention, Convention, and Utterance-Part Meaning
   - Introduction
   - A Nominalist Strategy for Part-Meaning and for Grammar
   - Occasional Part-Meaning and Occasional Grammar
   - Conventional Part-Meaning and Conventional Structure: Words and Grammar

3. A Gricean Theory of Reference
   - Intending to Refer
   - The 'Intend'-to-'Refer' Conditional
   - The 'Refer'-to-'Intend' Conditional
   - Referential/Attributive
   - The Epistemic Foundations

4. Referring: The Standard Theory
   - Identifying Knowledge
   - Searle on Reference
   - Strawson: A Problem About Grammatical Convention

5. Proper Names
   - Two Theories
   - The Standard Theory
   - The Causal Theory
5. Proper Names (continued)

4. Names and Referring 162

5. Application and Argument 171

6. A Strong and a Weak Version 184

Bibliographical References 191
Introduction

I shall be investigating referring, the typical function of proper names and definite descriptions. It has been said that one refers to an object when, in the course of uttering some sentence, one utters some term with the intention and for the purpose of picking that object out in order to say something about it. There is a historical and theoretical division of views concerning the preceding reference to speakers’ intentions: do they have a place in an explanation of reference, or can mention of a speaker’s intentions be limited to observations about his overall communication-intentions? A speaker very often makes some utterance for the purpose of communicating something to a hearer, and so we say he intends something by his utterance. It is at least initially an open question, though, whether he intends anything by the several parts of his utterance, and even whether his intention with regard to his utterance has anything essential to do with its meaning. Asking whether it is primarily speakers who refer or words which refer is a reflection of this difference in approach. One associates Frege, Russell, the early Wittgenstein, and the formally minded semanticians who follow Tarski with the latter position, and the later Wittgenstein, Grice, Austin, and Strawson with the former.

I shall be developing a theory of reference in which speakers' intentions with respect to the several parts of their utterances
have an essential place. Before I describe what I shall be doing in this essay, though, I want to set out the differences between a functional and intentional theory and its non-intentional opposite generated by what I call the syntactic approach. One begins an investigation of reference by, either tacitly or explicitly, distinguishing referring expressions from others. I think that there are at least two broad strategies one can follow: the distinction may be either a functional or a syntactic one. Leaving until later the question whether these are two truly alternative ways of distinguishing parts of utterances, I shall say a word about each.

If a kind of expression is distinguished syntactically, no appeal, or only a limited appeal, is made to what the speaker meant or intended. One looks, rather, to what we may call the formal or formalizable properties of utterances. Grammar studies these sorts of properties, and I have in mind particularly the sophisticated grammar we trace to Chomsky.¹ He uses certain complex criteria for judging a grammar adequate, and neither in the criteria nor in the theoretical terms of the grammar does one need, apparently, reference to speakers' intentions. A part of a sentence is of a particular grammatical category, e.g., distinguished as a subject

noun-phrase, if it has a certain role in the correct grammatical
description of the sentence, and particular arrangements of parts
which are of the appropriate grammatical categories make up gram­
matical sentences. Grammarians don't talk about referring expres­
sions, but one expects that they have or could construct a category
which contains all and only referring terms. The category of sub­
ject noun-phrase is too narrow to capture all referring terms,
but perhaps a combination of the categories of subject noun-phrase
and direct object might do.

Supposing that we can identify some part of an example sen­
tence as a subject noun-phrase, we have already described its role
in the sentence in functional terms, that is, it functions as a
subject. Grammarians use two sorts of distinctions, categorial
ones and functional ones. One distinguishes e.g. noun-phrases
and verbs categorically, and subjects and direct objects function­
ally. A categorical distinction tells in a non-relational way what
kind of thing a particular term is, while a functional distinction
marks what relations a given part has to other parts of the sentence.
The functional character of grammatical descriptions is, however,
a limited one, since any given grammatically distinguished part
has a (grammatical) role just with respect to the 'end' of making
the sentence grammatical.

Linguists such as Chomsky do not, so far as I know, commit
themselves on the question whether, in the order of explanation,
syntactic descriptions are logically prior to ones invoking
speakers' intentions. Many philosophers, however, think that the syntactic way of distinguishing parts of utterances is the right one. They believe that while an utterance without conventional internal structure may be said to have meaning, it may not be said also to have parts which have meaning. Being a meaningful part of an utterance, according to this view of the matter, is logically tied to being a conventionally distinguished part. I argue in detail against this view in Chapter 2, showing that an utterance may have meaningful parts with assignable grammatical functions, though those parts and functions are not distinguished conventionally.

When we characterize a referring expression functionally, we appeal to speakers' intentions, so the function which I shall oppose to the purely grammatical one is the speaker's intention in uttering an expression. He intends his utterance to serve a certain end, in a certain way. When a speaker produces an utterance with parts, he does so, in the non-conventional cases, by having distinguishing intentions with respect to the function of each part. We characterize a referring term functionally when we distinguish, on the grounds of its function, a part of the whole utterance as a referring expression. P.F. Strawson is the philosopher one associates with the functional theory of reference. His early work "On Referring" convinced many philosophers that, in the

first instance, referring is something people and not words do. People refer with a purpose, we presume; we give a functional explanation when we describe a speaker's intentions and how he intends to achieve them. Strawson's is not, in "On Referring", a purely functional characterization of reference; he apparently depends on grammar to pick out referring expressions.

I have purposely been describing the syntactic and functional strategies in quite general terms, since I wanted first to set out the two broad approaches, and later to indicate the nature of the theoretical tools I shall use to implement my program for distinguishing and describing referring and referring expressions.

I propose to develop what may fairly be called a Gricean theory of reference. It is H. P. Grice, of course, who I mean to invoke with the modifying adjective; his paper "Meaning"\(^3\) provides what I think is the essential tool and outline of a program for a theory of reference. Grice offers two distinct but fruitfully relatable theses about meaning; I shall apply them to reference.

A Gricean theory of reference will be a functional one, having those characteristics I mentioned above which distinguish functional explanations from syntactic ones. What Grice offers is (a) a formulation of a kind of speaker's intention which explains what it is for a speaker to mean something by an utterance in terms of what

he intends to achieve by that utterance and how he intends to achieve it, and (b) the outline of a program for relating what a speaker meant by a particular utterance on a particular occasion to what the utterance means. I shall use a generalized form of thesis (b) as a guide and a recommendation for relating any sort of convention-governed speech activity to its non-conventional analogue. The program suggested by thesis (b) gives the occasional, the non-conventional sorts of speech activity logical priority over the convention-governed sorts. It follows, according to this program, that a syntactic description, since it is essentially conventional, is somehow dependent on and subsidiary to a functional one.

Grice sees his program for meaning as giving a logical priority to occasional meaning, that is, to what a speaker meant by his utterance on a particular occasion. I shall discuss the relation between occasional and conventional meaning in Chapter 1. It has now been shown that the conventional meaning of an utterance is explainable in terms of occasional meanings, and in Chapter 2 I shall have something to say about how a similar program for grammar might go, at least in simple cases.

A Gricean theory of reference will parallel Grice's own theory of meaning in at least two respects. (a) Speakers' intentions are


primary: any explanation of how referring works begins with what I shall call the referring intention, that is, with what the speaker intends by his utterance of the referring part of his whole utterance. And, (b), the conventional reference of an expression is to be explained in some way which makes the occasional reference primary.

There is a third thesis which, together with the first two, characterize what I call a Gricean theory of reference. This last thesis serves to make clear the answer to a question I had postponed answering, namely whether the functional or syntactic characterization of referring expressions had logical priority. The third thesis is (c), that the conventional grammatical structure of an utterance, which distinguishes a part as, e.g., a subject noun-phrase and a referring term, is to be explained in terms of occasional grammatical structure. An utterance's ordinary grammatical structure is given by convention; it has an occasional structure just when a speaker distinguished parts in his utterance by intending them each to serve a unique function. This formulation does not, however, distinguish a speaker's intentions with regard to the meaning of the parts of his utterance from his intentions with regard to their grammatical function. Meaning and grammar are distinct and complementary, though: knowing the meanings of the words in an utterance isn't enough to give us the meaning of the utterance. In addition to knowing the meanings of the parts, we need to know how their
meanings combine, which is to say, we need to know what grammatical relations the parts bear to one another. Actually distinguishing these two aspects must wait until later.

Neither for meaning nor for grammar do I argue that the Gricean program is the right one. It is a way to proceed in relating non-conventional meaning to meaning, and it is a way of generating grammatical conventions. I don't know how to argue that it is the one right way, but I think that hardly matters, since in the meantime we can assess the program on the grounds of its explanatory power. The Gricean program has great power to explain and unify, as I hope I shall demonstrate. In particular, the questions of how the non-conventional use of an expression to mean something or to refer to something relates to the conventional use, and of how one can deviate, if one can, from grammatical convention, generate really difficult problems. These disappear, I shall argue, when one takes conventional uses to be related to non-conventional ones as the Gricean program relates them.

Chapters 1 and 2 are, with only a few exceptions, devoted to implementing the program of relating the occasional meaning and grammar of an utterance to the conventional meaning and grammar, and to exploring the consequences of so relating them. In the remainder of the essay, again with only a few exceptions, I deal with what, traditionally, are problems about reference. I hope to solve some of the standing problems about referring, and if with the use of a
generally Gricean theory I can, we have further evidence of the correctness and importance of Grice's original insight.

In Chapter 3, I formulate the conditions which a part of an utterance must meet if it is to refer. Also, I argue for three theses about reference. (a) If reference is to occur, there must be a causal connection between the object of reference and the speaker's state in referring. Someone who refers to an object must know that that object exists and, generally, somehow be in a position to pick that object out, and know himself to be able. I shall refer to the speaker's state in this respect as his epistemic connection with the object of reference. One can pick out an object in another way, which I describe as attributing a property to it, even if one lacks a causally based epistemic connection with that object. There are reasons for thinking this not to be a kind of referring, and it is anyway distinct from the causally based sort. And, the speaker's having identifying knowledge of a particular, that is, being able to recognize it, or knowing an individuating fact about it, is neither (b) necessary, nor (c) sufficient as an epistemic connection enabling him to refer to it.

Donnellan's\(^6\) work on reference provided me with important clues for all three theses about reference, and some form of a claim about

the importance of a causal connection appears in the work of Kaplan\textsuperscript{7} and Kripke\textsuperscript{8}. Arguments against requiring identifying knowledge of a referrer are found in, for the most part, Donnellan's paper "Proper Names and Identifying Descriptions", and in Kripke's "Naming and Necessity". I have stated conditions (b) and (c) negatively since in that way their relation to what I shall call the standard theory of reference is plainest. Strawson\textsuperscript{9} and Searle\textsuperscript{10} are the foremost of those who hold the standard theory. They both think that if one is to refer, one must have identifying knowledge of the object referred to. I shall present, in Chapters 3 and 4, the essentials of the standard theory, and shall try to disprove it.

In Chapter 5, I am concerned with the new theory of names which one associates with Donnellan, Kaplan, and Kripke, and which I call the causal theory. Strawson and Searle are again representative of those who defend the kind of theory of names I call standard. I


argue for the new theory in some detail, showing how it differs from the standard theory, and how it fits nicely with a Gricean theory of reference and indeed requires it.

Chapters 3, 4, and 5 contain the core of my discussion of what is traditionally considered theory of reference. Those Chapters, with the new and sometimes radically different answers they give to old problems, constitute what I describe as a Gricean and a causal theory of reference.
Chapter 1

Meaning and Convention

H. P. Grice's paper "Meaning" provides a solution to certain problems about meaning and also, I think, to problems about reference. His central and new idea has two parts. We can understand what it is for an utterance to have meaning by understanding what it is for a speaker to mean something in uttering it; and for a speaker to mean something by what he utters is for him to utter it with a certain kind of complex intention. In Chapter 2, I shall show how Grice's program extends naturally to parts of utterances. I apply his theory there in an attempt to generate meaningful parts of utterances without recourse to convention. In Chapter 3 I apply the theory to referring.

Grice offers an analysis of what he calls non-natural occasion meaning, or meaningnn. The meaning is occasional meaning because the analysis captures what a speaker meant, on a particular occasion, in or by making the utterance he did; the utterance's conventional meaning, if any, is inessential. In a natural sense of "mean", meaning_n, spots mean measles; we take the spots as a dependable natural indicator of measles. An utterance has non-natural meaning if a speaker means something by it in what I shall call the Gricean way. What is crucial is not whether the utterance x means_n some-

thing by x. If x has a natural meaning, a speaker who utters it (assuming it is an utterable kind of thing) has poorer grounds for expecting to be taken to mean something by x than if x had no natural meaning. His hearer will probably rest content at having noticed the natural meaning, supposing it to be appropriate to the occasion, and so the speaker's intention to mean something by x will be frustrated.

Grice concludes that:

"... we may say that 'A meant something by x' is roughly equivalent to 'A uttered x with the intention of inducing a belief (in the audience) by means of the recognition of this intention'."

The speaker intends in the Gricean way to produce a response in the hearer, and producing a belief is, according to Grice, what we want to do in simple cases of giving information. A speaker can have other purposes than the giving of information; commands, for instance, may be characterized by the speaker's intention to produce the effect of obedient behavior. The broader analysis will have the right-hand side of the equivalence as: 'A intends the utterance of x to produce some effect in an audience by means of the recognition of this intention'. I shall hereafter call this sort of reflexive intention 'the Gricean intention'.

As an essential feature of this analysis, Grice favors a pro-

gram we might call, borrowing the phrase from Bennett, meaning-nominalism. How should one order the following claims to place them in sequence of logical priority, beginning with the most fundamental? (i) A speaker means something by $x$, by or in uttering it with the appropriate intention; (ii) $x$ meant that so-and-so on a particular occasion, and what $x$ means is to be identified with what a speaker meant by it on that occasion; and, when $x$ is (almost) always, as a matter of rule or convention, used to mean a particular thing, we have (iii) $x$ means (timelessly) that so-and-so. A meaning-nominalism will order them as I have, moving from the more particular to the more general. Occasion meaning is fundamental, and conventional meaning is explained in terms of it. Meaning something by what one does is therefore something a person who was in other respects able could do even if he had no language whatever. Grice is a meaning-nominalist about whole utterances; he is not, as we shall see in Chapter 2, about parts of utterances.

I shall adopt a generalized meaning-nominalism as a program and a guide. I think that it correctly describes the relation between occasional and conventional meaning for whole utterances, in a way I shall shortly explain, and I shall attempt to extend this pattern of explanation to include conventional meanings of parts of utterances -- parts, when they have conventional meanings are words -- and of the conventions of grammar.

Grice's aims have been frustrated by a really oppressive number of counter-examples, almost all of which threaten to falsify just one half of his analysis. Since an equivalence is a pair of conditionals, in the analysis of meaning we have (1) if a speaker intends by his utterance to bring it about in the Gricean way that a hearer comes to believe that p, then the speaker meant that p by his utterance, and (2) if the speaker meant that p by his utterance, then he was intending by his utterance to bring it about in the Gricean way that the hearer comes to believe that p. The counter-examples have almost all been directed against (2), which I will call the 'mean'-to-'intend' conditional. Apparently not all cases of meaning something by what one says are cases of uttering something with the Gricean intention. Grice has chosen to defend the equivalence, but at a considerable cost. It is not clear that he is succeeding, and even if he is, the theory has become so fearfully complex as to weigh against its utility.

Bennett chooses to defend just the 'intend'-to-'mean' conditional, which the counter-examples in the literature don't threaten. He suggests that we give a partial analysis of meaning, demanding only sufficient conditions on something's being a case of meaning. The unanswered difficulties about the converse conditional will have to be faced sometime, but if the partial analysis captures a

central enough aspect of meaning, we stand to gain considerably in understanding. And, if the aspect of meaning we can explain in the Gricean way has the right properties, it may serve as the base upon which to build an explanation of all sorts of meaning. I shall show that, while no complete explanation is in sight, conventional meaning, which generates most of the counter-examples to the 'mean'-to-'intend' conditional, can be explained in terms of the basic kind of Gricean meaning we do understand.

Searle's attempts a vaguely Gricean analysis of referring. I shall argue that he doesn't succeed because, in part, he doesn't appreciate the virtues of the original Gricean analysis of meaning. Searle replaces Grice's analysis of meaning with one only slightly related to the original. I shall spend a few pages arguing against Searle's view of meaning for two reasons: the errors in his theory of meaning cause errors in his theory of reference, and his views on conventional meaning (though not necessarily his particular way with them) are, I think, widely held.

Searle thinks that making an utterance and meaning something by it does not connect in any essential way with a speaker's having the intention to induce a belief in his hearer. He objects that Grice's analysis links meaning with the production, Searle thinks,

of the wrong kinds of effects in the hearer. He says:

"The characteristic intended effect of meaning is understanding, but understanding is not the sort of effect that is included in Grice's examples of effects. It is not a perlocutionary effect. Nor can we amend Grice's account so that meaning is analyzed in terms of understanding. That would be too circular, for one feels that meaning and understanding are too closely tied for the latter to be the basis for an analysis of the former."  

Notice that this objection does not bear directly on the 'intend'-to-'mean' conditional. Searle says that when one says something and means it, one doesn't necessarily intend to produce the effect of belief in a hearer, but may intend simply to produce understanding. But Searle doesn't tell us whether he thinks that a speaker's intending to produce a belief in the Gricean way can ever be a case of meaning. That much of what Searle has to say about the matter seems right. Grice has, since Searle wrote what I have quoted, said that 'come to have the thought that' may be used instead of 'come to believe that'. This formulation of the intended effect avoids some of the difficulties Searle finds. In particular, it allows for cases in which the hearer comes to know what the speaker is trying to get him to believe, while reserving judgment on it.

From this objection about intended effects, Searle moves rather too quickly to a point about conventional meanings; he appears not to notice Grice's explicitly mentioned hope that conventional meaning may be explained in terms of occasional meaning.

Since Searle ignores that point, he brings as an objection to Grice's account that it doesn't show how "meaning is more than a matter of intention, it is at least sometimes a matter of convention." Grice didn't show how meaning was sometimes a matter of convention, but he certainly doesn't claim that it is not. But Searle seems to claim that meaning is always a matter of convention. He says that a speaker utters some sentence T and means it if and only if he intends in the Gricean reflexive way that his hearer come to understand what he says, and he intends that this intention "will be recognized in virtue of (by means of) H's knowledge of (certain of) the rules governing the elements of T". We should probably construe Searle generously, applying his analysis of meaning only to cases in which one can utter T and "mean it"; this restricts Searle's analysis to cases in which T has a conventional meaning. We are, on this understanding of what he is doing, free to supplement Searle's account of meaning with the 'intend'-to-'mean' conditional, claiming that it is appropriate to cases of meaning in which no convention operates.

Searle might find this two-part theory of meaning congenial, but it won't do; there will be severe and I think crippling difficulties with deviations from convention. Given Searle's account of what it is for a speaker to mean something by an utterance which has conventional meaning, how can a speaker, ever, mean something

8. Searle, p. 50.
by the words he utters other than what they conventionally mean?

Searle does see that this is a problem.

"In the Philosophical Investigations, Wittgenstein (discussing a different problem) writes "Say 'it's cold here' and mean 'it's warm here'". The reason we are unable to do this without further stage setting is that what we can mean is at least sometimes a function of what we are saying. Meaning is more than a matter of intention, it is also at least sometimes a matter of convention."

This passage implies that we sometimes can mean something by what we say other than what the words we say mean. Searle doesn't follow it up, though; he nowhere explains how we can deviate from convention. I suspect that he doesn't because the shape of his theory forbids it. If the characteristic intended effect of meaning is understanding, how can a speaker ever intend his hearer to understand him to mean by a word something other than what it conventionally means? Searle specifically says that when I use a sentence with a conventional meaning, I must, if I am to mean anything by it, intend my hearer to understand it to mean what it conventionally does.

There is almost no conceptual space left in which to maneuver -- but there is a little. One can have various reasons for deviating from a convention. Suppose that one is speaking to someone who inverts the meanings of 'hot' and 'cold'. One could appeal to his false beliefs about conventional meanings to explain how, in speaking to him, one can deviate from convention by oneself purposely.

inverting their meanings: one speaks as if they had the conventional meanings one's hearer thinks them to have, one speaks according to his (understanding of the) convention. This only works for one kind of deviation from convention, but it seems true that a speaker can use words deviantly whenever he expects that his hearer will understand what he means by them. Searle might object to this claim, since it is an element of a Gricean-type theory of meaning and convention. I shall therefore not assume that deviation from convention is explained by the above theoretical claim (though I think it is) but shall simply adduce deviation as a fact. I can use words deviantly in cases where my hearer has no false beliefs about the conventional means of the words I use, and sometimes I will be understood. It seems to me that the best Searle can do with this fact is to claim that sometimes we pretend that, or speak as if, a word has a conventional meaning other than the one it has. If our hearer happens to figure out what the pretend-convention is, he understands what we mean.

I wonder, though, just what force is left to convention if we can, supposing our hearer to be acute enough, operate with pretend-conventions just as well as with real ones. One wishes that Searle were more explicit about what sort of thing a convention is; if he were, we might be able to tell if they can sometimes work in this way.

As against this rather messy situation, I offer the simple,
satisfying, and theoretically powerful combination of Grice's theory of meaning and Lewis's theory of convention. Bennett shows us the strength of such a combination; his paper "The Meaning-Nominalist Strategy" gives the outlines of a unified theory of meaning.

The Grice-Lewis program will give us conventional meaning, will make conventional meaning a species of the genus meaning, and will explain how we can intend, on occasion, to mean something by our words other than what they conventionally mean. Bennett takes an analysis of convention from Lewis and broadens it, suit ing it for application to the speaker-hearer relationship. This analysis has the properties needed: it is enough stronger, for the right reasons, than a regularity, and it builds on occasion meaning.

Lewis' analysis of convention, in summary, is as follows. People need to co-ordinate their actions when it arises that what it would be best for any one of them to do depends on what the others do. Suppose we are speaking on the telephone and our connection is somehow broken. Each of us wants to re-dial if and only if the other hangs up and waits — our problem is to co-ordinate our actions. A co-ordination equilibrium for a group results when it is true of each member $x$, that given what all the others do, there is no action which $x$ could have performed which would have given results he would have preferred to those resulting

from the action he did perform. Co-ordination equilibria are
important because when one is achieved, no member of the group has
reason to regret the way he acted, given how all the others acted.
A co-ordination equilibrium is one of a set of solutions to a co-
ordination problem, and leads to maximising the benefits accruing
to each person involved with respect to the problem requiring co-
ordination, given the restraints of the situation, namely the
actions of the others involved.

Co-ordination may regularly be achieved because of a prior
arrangement, or for some other reason. A convention, in Lewis'
sense, is a case of a co-ordination equilibrium which is effected
or maintained in new situations for certain kinds of reasons.
There is a convention if the following obtains. A group faces a
recurring co-ordination problem whenever they are in a situation
of kind K. They have regularly in the past resolved this problem
by acting in a certain way, say using procedure C, and it doesn't
matter how this regularity started. It is mutual knowledge, as I
will call it, within the group that C has been followed in the past.
This means that each member of the group knows, and knows that the
others know, and knows that they know that he knows, and so on,
that in K situations procedure C has been followed usually or reg-
ularly in the past. In a new K situation, if the group cannot con-
sult on how to act, they are very likely to follow C once again.
It is the only procedure, barring some unusual occurrence which
causes them to think otherwise, that each can reasonably believe likely to be the one the others will follow. So, if any member of the group believes that C is the procedure the others are most likely to follow, it is reasonable for him too to follow C, and so it will be reasonable for each to follow C.

The behavioral regularity -- in K situations procedure C is followed -- is maintained in the new K situation because it is mutually known to have been followed in the past; each member of the group has a reason for doing C again -- it is the way he should act to maximize the likelihood of success. A regularity of this sort maintained for these reasons is a convention.

Bennett wants to apply Lewis' analysis of convention to meaning. To do this he has to find what is being co-ordinated when a speaker speaks and a hearer understands. Lewis analyzes the co-ordination of actions properly so-called, but a hearer's understanding of what is said is not an action. Lewis applies his analysis to a kind of communication which he calls 'signaling', in which understanding is associated with action in such a way that an action a speaker performs in signaling is co-ordinated with an action performed by the hearer. The hearer's response may be a signal, or some other action, but because there must be some action on the hearer's part which is co-ordinated with the speaker's signal, we've left out most ordinary cases of communication conventions.

When a speaker speaks and a hearer understands but does not
say or do anything in reply, what the hearer understands (the speaker to mean) is sometimes governed by the conventional meaning of the expression used. If we cannot co-ordinate what the speaker says with what the hearer understands, we will not be able to include that case as one governed by convention, though it clearly is. We could attempt to co-ordinate what the speaker says with what the hearer does or would reply; but a reply or some other action is just not always appropriate, nor do we want to be forced to include a decision as to what is the correct reply in our solution to the problem.

Bennett solves the problem by extending Lewis' analysis of convention, so that it applies to what Bennett (on a suggestion from the present writer) calls "doings". Actions will be included as a species of doing, though other things as well may be a doing and be governed by convention, e.g. coming to believe something. A person does something if there is some \( \theta \) such that the person \( \theta \)s at time \( t \), and one's beliefs can be reasons for one's \( \theta \)ing. All human voluntary actions will therefore be doings, and so too will be meaning something by what one utters, and coming to believe something. Intendings and belief-acquisitions are doings, and so is the acquisition of any state with a belief-component, such as fear, hope, etc. Doings are a very broad class of item, but their utility for a theory of meaning lies in their belief-related aspect. Doings group together actions (for which we can have motives other than beliefs
and which are under voluntary control) and e.g. belief-acquisitions
(which are not voluntary, but for which we have reasons). The con­
necting element is that for both, we can have reasons for doing them.
Lewis' analysis of convention can be rewritten with 'action' replaced
throughout by 'doing'. It will work because the only motives for
action relevant to co-ordination problems are beliefs, plus a desire
for whatever it is that requires the solution of the co-ordination
problem.

Bennett wants to apply the broadened concept of convention to
a speaker S and a hearer H. S wants to get H to believe that p by
uttering U, intending in the Gricean way that U mean that p. H,
for his part, wants to know why S does U; he wonders "What is he
trying to tell me?". Co-ordination is achieved if S utters U with
the Gricean intention to mean that p, and H recognizes S's intention,
and so comes to believe or have the thought that p. If S and H
mutually know that, in the past, whenever someone in their language-
group uttered U he meant that p, and was so understood, S and H each
have a good reason to do the same again. S utters U and means by
it that p, and H comes to believe or have the thought that p.

An utterance U comes to have a conventional meaning p when mem-
bers of the group adhere to the regularity that by uttering U one
means that p because that is the way of acting most likely to achieve
their ends; it is the best way, owing to their mutual knowledge that
the regularity connecting U and p has previously held. We can now
generate conventional whole utterance meanings, but we are still some distance from conventional word meanings. One really serious difficulty has been removed, though.

A convention about meanings will provide grounds for saying, among other things, that deviant utterances are not just deviant but, usually, wrong. We can sometimes legitimately flout a convention. How? Donnellan\textsuperscript{11} has offered a really satisfactory explanation along Gricean lines of why one cannot intend to refer, using some expression, to just any object one might pick. The (false) presumption that one can arises from defining reference as something like: "intending to single out some object for one's hearer, and...". However the explanation is filled in, if one fails to give the expression one uses an essential role in fixing the referent, the presumption may be given that the conventional reference of the expression one uses has no role at all in picking out the referent. Donnellan succeeds in showing both how the conventional reference of an expression matters, and how sometimes one can and sometimes one cannot refer deviantly. I'm concerned with deviance from referring conventions as a special case of deviance from linguistic conventions. The three sorts of linguistic convention are those for reference, meaning, and grammar. In each case the possibility of

convention-flouting as well as the limitations on that possibility come from the connection between what one can intend and what one can expect to succeed at doing.

If one intends to refer to some particular object, one intends to single it out for one's hearer. One singles out that object for one's hearer if and only if he comes to know which object it is one intends to single out, and that one did intend to single it out for him. It is a conceptual truth about intending that one cannot intend to do something at which one expects to fail, and so one cannot intend to refer unless one expects to succeed at it. What expectations one has will be tied to the beliefs one has, and this is a logical point. For instance, I think it highly unlikely that anyone believes that the Prime Minister of the U.K. is the King of England, and so, in ordinary circumstances I just cannot intend to refer to the Prime Minister of the U.K. using the expression 'the King of England' since my intention would in all likelihood be frustrated.

The argument works, explicitly, through this conditional: if one refers to something by what one utters to a hearer, then one intends to bring it about that the hearer comes to know which object is being referred to. If the consequent cannot be realized because one cannot expect to succeed at realizing one's intention, then the antecedent will be false, and one does not refer. The presence of a convention limits what one can intend because it makes some
intentions unrealizable, as for instance my 'intention' to refer to the Prime Minister of the U.K. using 'the King of England'. I can refer (in this central way) only if I expect my audience will come to know, through what I do, which object I want to single out. My knowledge of the situation determines or should determine what I expect to succeed at. Only in certain situations can I use a particular expression to refer to a particular object because in only certain situations can I expect to succeed at picking it out for a hearer. I am not free to intend anything I like by my words; what I can intend is tied to what I can expect to succeed at.

When a speaker uses a referring expression which has a conventional referent, his hearer in knowing the convention has a standing expectation that the expression will be used to single out its conventional referent. The usefulness of conventions about meaning and reference lies in just this, that the hearer supposes, unless he has reason to believe otherwise, that what the expression is used to mean or refer to is what it conventionally means or refers to. Hearers aren't always in a position to make complicated inferences to what a speaker intends; but when a hearer isn't, a speaker can depend on conventions about meaning and reference, a hearer can so understand him, and communication is achieved. Most of the time the conventional aspects of referent-determining will dominate the field. They do not when, for whatever reason, a speaker thinks he can override or somehow circumvent the hearer's standing expectation.
that the expression used is meant conventionally.

Donnellan's account applies to meaning as well as to reference. Generally, one can intend by some expression whatever one can reasonably expect an audience to take one to mean by it; and in particular, one can mean something by the expression other than its conventional meaning only if one reasonably expects to be understood. Wittgenstein's challenge to "say 'it's cold here' and mean 'it's hot'" will do as an example. I can say 'it's cold' and mean 'it's hot' just when I can expect to be understood as meaning 'hot' by 'cold'. I could expect this if, for instance, I knew I was speaking to someone who had inverted the meanings of 'hot' and 'cold'. Because of my hearer's false beliefs about the conventional meanings of those words, I can use them deviantly, confident of being understood.

This argument, like the one about referring, uses a form of the 'mean'-to-'intend' conditional, namely, "if one means something by what one utters to a hearer, then one intends, via the Gricean intention, to bring one's hearer to believe something". This conditional is perhaps not as obviously false as the strong, original one, but an even weaker one will serve, which seems unlikely to be questioned. We can use "if someone means something by what he utters to a hearer, he intends it to have a certain effect on the hearer", so long as what he means depends on what he intends. The consequent is falsified when the speaker cannot have the appropriate intention, and
this entails the falsity of the antecedent. We conclude that the speaker did not mean anything by his utterance because he did not have the appropriate intention.

From the observation that what a speaker means depends on what he intends (and what he intends on what he expects he can achieve), we see that the place of convention is just to bring it about that one can expect to succeed at conveying to a hearer that p. If the hearer knows the convention governing a certain expression U, he has a very good reason to believe that the speaker who utters U intends to get him to believe that p because that is what U conventionally means. In a particular case, one can mean something by an utterance other than what it conventionally means just when one believes that one can overcome one's hearer's standing expectation that the expression, in this instance, means what it conventionally does. Because this situation will relatively seldom arise, the conventional meaning of an expression is by far the most important restriction on what one can mean by it.

Taken by itself, one might have reason to be at least somewhat sceptical of a Donnellan-type treatment of convention flouting. Unless one knows what a convention is, one can't be sure that one can be overridden in the way he tells us it can. Lewis' account of convention removes the problem. A convention is, according to Lewis, the sort of thing which has force just when one's hearer expects that one is not deviating from it. He tells us that a
convention is one kind of solution to a co-ordination problem; but
the problem exists independently of the convention, and other solu-
tions are possible. A convention can be flouted when some other
solution is available, and cannot be just when, for circumstantial
reasons, the conventional solution is the only reasonable one. So,
for instance when Donnellan gives as an example of convention flout-
ing a story about a usurper who is called 'the King' by everyone,
though everyone knows he is not the King, we can see easily enough
why convention can be flouted. The usurper demands to be called
'the King', and so everyone knows who everyone else is referring
to when they refer using 'the King'.
Chapter 2

Intention, Convention, and Utterance-Part Meaning

1. Introduction

There are two theses about the meanings of parts of utterances which I think are widely held among philosophers: (1) that a part of some utterance is meaningful only if instances of that part-type do or would (in circumstances to be specified) appear in various utterances, contributing each time the same or much the same to the meaning of the whole; and (2) that the meaning of any part is, in some way yet to be explained, derived from a convention, rule, or regularity governing uses of the instanced part-type. I shall call the first the recurrence thesis and the second the convention thesis. The first I argue is true, and the second false. These two theses relate to the distinction between words and meaningful parts. A word is a meaningful part which satisfies (2); words have conventional meanings, so it is always appropriate to ask what a word means, and not just what a speaker meant by an utterance containing it. A part, by contrast, is meaningful if it satisfied (1); if (2) is false, the part may be meaningful just because a speaker means something by it, independently of any role or convention.

The contrast between (1) and (2) above for parts parallels the contrast for whole utterances between occasional meaning (what a speaker meant on an occasion of utterance) and the conventional
meaning (what the utterance-type means). Bennett\(^1\) has suggested that "meaning-nominalism" be used to stand for the following two theses about meaning: (1) that an utterance can be meaningful without having a conventional meaning, and (2) that the conventional meaning of an utterance-type is to be explained in terms of the occasional meaning it has had on past occasions of use. The two theses may be extended to include parts of utterances and grammatical structure: (1) a part can be meaningful without having a conventional meaning, and an utterance can have grammatical structure independently of convention; and (2) the conventional meanings of part-types and the conventional structures of utterances are to be explained in terms of occasional meaning and occasional grammatical structure. Successfully applying meaning-nominalism to parts will demonstrate the falsity of the convention thesis.

Meaning-nominalism is a Gricean program: in his original paper "Meaning"\(^2\), Grice suggests that we may be able to explain the conventional or timeless meaning of some utterance-type in terms of the meanings it has had on past occasions of use. Bennett attempts to implement this suggestion; his program is examined briefly in Section 2. That program can be extended to parts of utterances, but in doing


so, will go counter to Grice's own preferred way for advancing from whole utterances to parts: he is not nominalist about parts, but holds a version of the convention thesis in which 'regularity' replaces 'convention'. I shall argue directly against Grice by means of counter-examples embedded in a theory which I think explains how a part of an utterance may be meaningful on an occasion, without invoking any rule, regularity, or convention. I shall argue that, just as with occasional and conventional whole-utterance meaning, meaningful parts are logically prior to conventionally meaningful parts, and that the former may be used to explain the latter and not vice versa. The development of an extended meaning nominalism occurs in Sections 3 and 4, and in Section 5 Lewis' analysis of convention is used to generate conventional part-meanings and conventional grammatical structure.

Those who believe both the recurrence and the convention theses about part-meaning have roughly this picture of the logical genesis of parts with meaning. Some group of communicators engage in discourse first using utterances which lack meaningful parts. Some of those utterances acquire conventional meanings, and in some way yet to be explained, internal structure. From the set of utterances


with conventional meanings, one picks out occurrences of some part-type by physical characteristics, finding that they contribute the same to the meaning of each of the set of utterances. Such a part is meaningful, and it arises from and is distinguished through regularities, rules, or conventions.

The convention thesis says that part-meanings are distinguished through, and come into existence by means of, regularities or rules holding among some set of whole utterances. It is hard to find specific arguments advanced by particular philosophers in defence of the claim, and while there are some plausible arguments, I think they ultimately fail.

Frege\(^5\) warns us not to ask about the meaning of a word except in the context of sentences in which it occurs. He offers this advice as one of his three maxims for avoiding what he calls psychologism about meanings. We are not interested, he rightly says, in what sorts of associations people have with words, but in what the words mean. He says that what words mean are in important respects speaker-independent, and couples this with the claim that we should understand the meanings of words in terms of the contributions they make to the meanings and truth-conditions of the sentences in which they occur. Frege's observations were in part responsible for the move away from theories of meaning in which a word has meaning

because it 'stands for' an idea or 'expresses' a concept a speaker has. Such theories are of course wrong, and to see that words or meaningful parts cannot be logically prior to whole utterances is to see something important about why they are wrong.

What Frege saw was that the meanings of words cannot be investigated except by investigating the contributions they make to the meanings of whole utterances: whole utterances are logically prior to meaningful parts. Frege was right, but there is a certain crucial unclarity here. There is first the question of relative independence: there can be whole utterances without meaningful parts, but there cannot be meaningful parts unless those parts occur in and contribute to the meanings of whole utterances. The difficulty comes in sight when we ask in what way investigations of the meanings of parts relate to investigations of the meanings of wholes. It is clear that we must explain the meanings of parts in terms of the contributions they make to the meanings of wholes, and from this, it seems that there is an epistemic priority between the two: that we must know the meaning of wholes before we can know the meanings of any parts. Part-meaning will, according to this program, be explained by elucidating the systematic contributions some part-type makes to the meanings of the whole utterances in which it occurs. The epistemic priority claimed is this: in order to discover the meaning of some part, we must first know the meanings of one or several whole utterances in which it occurs.
I deny that there is in fact an order of epistemological priority. We may retain the thesis that a part is meaningful only insofar as it contributes to the meaning of some whole utterance, while discarding the thesis that we must know the meaning of that whole utterance (or of some whole utterances containing the part) in order to know the meanings of the parts it contains. This way is open to us, and one reason for the almost universal acceptance of the epistemic aspects of the whole-before-part view is simple oversight. But in addition, there is a certain difficulty with denying the whole-before-part view, a difficulty about criteria for identifying meaningful parts and determining their meanings. What kinds of reasons, what kinds of evidence, can we have for distributing the meaning of the whole over the parts in one way rather than another? The evidence we have for picking out whole utterances and assigning meanings to them doesn't in any obvious way extend to parts. In the absence of such evidence, it is natural that we should think that just how the meaning is distributed over the parts is a matter of convention. What I deny, of course, is that in the absence of convention, the way in which we distribute meaning over the parts of an utterance is necessarily arbitrary. I shall argue that in logically primitive cases of utterances with structure, we may have adequate evidence of a behavioral sort that some particular assignment of meanings to parts is correct.
2. A Nominalist Strategy for Part-Meaning and for Grammar

Bennett deploys Grice's analysis of meaning to show how an utterance can have meaning on an occasion, without benefit of convention, and brings a revised version of Lewis' analysis of convention to show how the conventional meaning of an utterance-type can be explained in terms of the meanings it has had on past occasions of use. I shall proceed in the same way in generating meaningful parts. Bennett only gives sufficient conditions for an utterance's meaningfulness; I think that difficulties with necessary conditions in Gricean analyses relate critically to convention. We sometimes speak in conventional ways because we know them to be correct; and we almost always depend on the conventions to provide our utterings with internal structure. And, once we suppose conventions are operating, we look to them, rather than to speakers' intentions, in our investigation of meanings. Speakers' intentions aren't relevant to an investigation of the meanings of words in a language; non-conventionally meaningful parts, by contrast, are non-linguistic because non-conventional.

I shall defend two theses of an extended meaning-nominalism, (1) that a meaningful whole utterance may have meaningful parts and grammatical structure entirely without benefit of convention, rule, or regularity, and (2) that conventional part-meaning (word-meaning) and conventional grammatical structure are to be explained in terms of their non-conventional counterparts. The convention thesis about
meaningful parts will have been shown to be false if (1) and (2) above are shown to be true. But while the convention thesis is I think false, the recurrence thesis is true, and captures something important about occasional part-meaning: if a part P of some speaker's utterance is meaningful for him, he does or would (in circumstances to be specified) use it again in other utterances in a way such that it contributes much the same to the meanings of each of the containing utterances.

The particular thesis about part-meaning I shall be concerned with is this: P is a meaningful part of U for S if in uttering P in the course of uttering U, S intends in the Gricean way that P be effective in a certain way in realizing S's Gricean intention for the whole utterance U. Unless P is a proper part of U, of course, the case is degenerate and requires no special treatment. A part is meaningful, I shall argue, if it contributes in a certain way to the realization of the speaker's overall meaning-intention; a certain aspect of the speaker's intention for his whole utterance concerns the function of some part, and other aspects concern the roles of other parts. Since the mode of contribution is that of a Gricean intention, we ensure that what we have is a meaning-like property of the part: the content of S's intention for P, that is to say, the function he intends it to have, tells us what P contributes to the grammar and meaning of U, and so what P's meaning and grammatical role is.
There are two kinds of demand a nominalist program for parts must be able to meet. First, a generally Gricean and intentional characterization of meaningful parts must be able to show them to be really distinct, with particular meanings, and second, it must be able to give a precise account of what a speaker intends when he means to give his utterance some certain structure and the parts certain meanings. I shall begin with the first, the sceptical ques­tion. In Section 3 I construct the Gricean intention which lets a speaker intend his utterance to be of a referring-and-predicating structure, since for reasons which will appear, that structure is both fundamental and peculiarly apt as a first case. For very com­pelling practical reasons, concerning the degree of complexity in intention a speaker is able to have and a hearer to recognize, I introduce next a conventional referring-and-predicating structure, using intentions operating on this simple conventional structure to generate more complex grammatical forms. At some upper level of complexity, structure cannot be achieved by intention, but pre­cisely when convention becomes dominant and intention subsidiary isn't important.

One aspect of the distinctness of a meaningful part, and one criterion for its distinctness, is its essential role in making possible the realization of the speaker's intentions for his whole utterance. In logically primitive cases, we are not justified in claiming that an utterance has parts unless the speaker expects
that his hearer will recognize what the parts contribute to the whole; and generally this will occur only when hearers' recognition of parts is essential to hearers' recognition of the meaning of the whole. We can weaken the condition slightly, though, allowing that if a speaker's expectation of success in communicating that \( p \) is increased by having intentions for parts of his utterance as aspects of his intention for the whole, he may distinguish parts. The point is, that if recognition of the meanings of parts would in no way contribute to the likelihood of recognition of the meaning of the whole, the speaker can't expect his intentions for the parts to be recognized, and so he can't have such intentions.

Not worrying about the precise formulation of the speaker's intention, we should have this as the speaker-hearer situation in a first case of an uttering's having a referring-and-predicating structure. The speaker believes that giving his utterance structure is a good way to achieve his communication-aim, namely inducing in the Gricean way the belief that, e.g., \( X \) is \( \emptyset \), and so his intention for the whole has part-directed aspects. The hearer wonders why the speaker says what he does, and concludes (1) that the speaker is trying to get him to believe something, (2) that the utterance has some sort of internal structure, (3) that the first part has the function of picking out the object of the belief, and that \( X \) is that object, (4) that the second part has the function of attributing some property \( \emptyset \) to the object picked out, and (5) that the
utterance means that the object picked out by the first part has the property attributed to it by the second, that is, that X is \( \emptyset \).

For instance, the referring part might be a pointing gesture, and the predicate-part a distorted-hand-presenting gesture, the whole utterance meaning "that wolf has a crippled foot". I discuss this example in detail in the next Section, but for the moment I wish merely to illustrate how a part may contribute to the meaning of the whole: the pointing picks out a thing, and the distorted-hand gesture, by virtue of its iconic element, suggests a property.

Later I shall be attempting to show in what circumstances a speaker could expect to convey to a hearer "that wolf has a crippled foot" by means of the hearer's recognition of the relation between the two gestures and the two aspects of the speaker's intention.

I have said that a speaker's meaning-intention is always about a whole utterance, though perhaps with aspects which relate to the functions of the parts of the utterance. It is clear from the example why there must be just one intention with two aspects, rather than two distinct intentions: for what could they be? How could a speaker intend his hearer to come to believe "that wolf" or "has a crippled foot"? We could, however, talk of two intentions, so long as we realize that they are not really logically distinct, since they cannot occur independently of one another. So doing will in fact prove the simplest and clearest way to display the structural relations between a subject-term and a predicate-term.
Notice that the aspect of the speaker's Gricean intention which characterizes a meaningful part involves reference only to part of what the speaker intends by the containing whole utterance. In a particular case it might come to: a speaker intends in the Gricean way that a part \( P \) be effective in some way, e.g. by referring, as an aspect of his intention to get his hearer to believe something. Unless the content of the belief is properly expressed in subject-predicate form, of course, referring will be out of place, so the ways we characterize the part depend on the intention for the whole. But, how we characterize the part doesn't depend on what the content of the belief is: we needn't know the meaning of the whole. Had it, we would not be able to distinguish parts except in a way which essentially involves knowledge of the meaning of the whole; and we would not, therefore, be able to identify a thing as a meaningful part of a whole utterance and determine its meaning without having independently determined the meaning of the whole. Had our explanation of part-meaning presumed such knowledge, the program would no longer be a fully nominalist one. If we, and the hearer, had to know the meaning of the whole in order to know the meaning of a part, recognition of the part-involving aspects of the speaker's intention must have been inessential to learning the meaning of the whole utterance. The speaker could not then have had an intention with part-directed aspects since (1) they would be inessential -- we can tell the meaning of the whole perfectly well
without worrying about the meanings of parts, and since (2) the speaker cannot expect us to recognize part-directed aspects of his intention because they are, by (1) above, purely gratuitous. One concludes, following this line of argument, that there is no Gricean and intentional way to generate parts if we must know the meaning of the whole utterance in order to know what the parts mean. This route leads to the standard theory in which meaningful parts are derived from regularities in meaningful whole utterances.

Sometimes a speaker cannot hope to communicate that p in the Gricean way unless he intends his utterance to have internal structure; and sometimes so intending increases the likelihood of his succeeding. This is one aspect of the distinctness of meaningful parts in the primitive case. The other aspect consists in facts about the speaker.

An utterance has meaningful parts in the primitive case because a speaker has a Gricean intention with aspects directed at each of several physically distinct parts of his utterance. He could use any of those parts again, in various combinations, to make new whole utterances, and he could use a part again so that it adds the same to the meaning of some new utterance as it added to the meaning of the old. For instance, the speaker in the wolf-hunting example could use the pointing gesture in combination with various predicate-terms, in a way such that the pointing makes a certain systematic contribution to the meaning of each utterance in which
it occurs: it picks out an object. The distorted-hand gesture also could be used in various utterances, though it is less easy to imagine a variety of cases in which a term which means "has a distorted hand (or foot)" might be useful.

A speaker in this sense has a part he has used previously available for use whenever he might find it suitable. It would be too severe to require that he would use it again whenever circumstances indicated. He may not remember how or that he has used it in the past, and he need not use it in that same way again. But a speaker has some part available for use on suitable occasions in this weaker sense, that if he remembers having used it in the past, (1) he can use it again if he wishes, combined with a new predicate-term, such that it serves in the new utterance the same or much the same function as in the old, and (2) he would very probably use it again if it would serve his purposes at least as well as some other expression. So the recurrence thesis is true: if a part of some utterance is meaningful, it does or would, under the specified circumstances, appear in various utterances, contributing much the same on each occasion to the meaning of the whole.

The whole-before-part program is wrong precisely because, in the absence of conventionally meaningful parts, we don't need to know the meaning of the whole utterance in order to know the meanings of its parts. Someone can make an utterance in the fundamental Gricean way and I can conclude that the speaker meant to effect
me in some way by his utterance, that he intended to get me to believe something, that the first part of his two-part utterance was somehow about the object X, that it was intended to single out X for me as the object of the belief the speaker intends to convey, and as the object to which the second part of the utterance attributes some property. I have learned the meaning of the first part, and I could learn the meaning of the second, without previously learning and then using the meaning of the whole. We have an explanation of what it is to be a meaningful part which neither in principle nor in practice presupposes that we know whole utterance meanings.

When a speaker has the right sort of intention regarding some part, he thereby distinguishes it as a meaningful part. He intends, in the Gricean way, that it have a certain function; and the function he gives it constitutes its grammatical role as well as its meaning. Meaningful parts are marked out by aspects of a speaker's meaning-intention, and the functional relations between those parts, as given by the relations between the aspects of the intention, provide the grammatical structure of the utterance. In a subject-predicate utterance, for instance, the structure is simple. There is a referring and a predicating part; the referring part has the grammatical role of subject-term for the predicate-part. The more meaning-like property of the referring term is simply that it picks out the object that it does; the predicate's meaning is related to
the property it is used to attribute.

Just as we did with meaning, we can investigate the occasional structure before we investigate the timeless or conventional structure. We have the conceptual means to assign grammatical structure in the absence of convention, and so as with meaning, we can go from the occasional structure, via Lewis' analysis of convention, to the conventional structure; occasional structure is logically prior to conventional structure.

3. Occasional Part-Meaning and Occasional Grammar

I shall construct some imaginary speech situations which might arise for a community of people who wish to communicate but have no language; the examples are designed to show how the group might begin producing and using utterances with internal structure. It is easiest to suppose that they already can communicate with whole utterances in the Gricean way, though this is not necessary to generating utterances with structure. It seems in principle possible that the first meaningful whole utterance produced in the group should have semantic structure, though the practical reasons against this occurring are overwhelming. The speech situations and utterances are presented in an order of increasing grammatical richness because complex grammatical structures develop most easily if a speaker can depend on his hearer having some previous experience at using utterances with parts. The constructions will of course be incomplete, since I neglect that which doesn't bear directly; how
the sketches could be filled in should be fairly obvious.

While I shall be examining how more complex communication-means develop from simpler ones, what is important is what may be learned in this way about the concepts fundamental to language. For that reason, we may remain agnostic as to how language in fact arose, while finding an imaginative reconstruction of language from logically primitive beginnings enlightening.

Grice's analysis of meaning tells us that in the simplest sort of fact-stating discourse, a speaker intends to induce a belief in his hearer, and what he means is the content of that belief. When the belief is 'about' an object, it may be rendered in either subject-predicate or existential form, and it is not obvious how we are to choose between them. Sometimes we cannot, but sometimes, I shall argue, we have in speaker-hearer situations evidence purely behavioral in kind which allows us to conclude that a subject-predicate rendering of the speaker's meaning is correct. I do not argue that all statements which are in some sense 'about' particular objects can be classified, on purely behavioral grounds, as subject-predicate or as existential: there remains a core of statements which I think refuse classification. But at the same time, there is something to be said about the general case; there are some considerations, some guidelines, which we may use to sort primitive statements without a conventional grammatical structure into subject-predicate, existential, and neither.
Consider a case in which someone comes back to the tribal village, and conveys in an iconic way that (the) bananas are ripe, by miming to his hearer the picking, peeling, and eating of a banana. Suppose that the bananas have just begun to come ripe, so the hearer has no reason to suppose that all the bananas in their area have come ripe, but that some have. The choice of renderings is first between (1) "the bananas are ripe" (in some area, on some certain tree, or trees) and (2) "there are some ripe bananas". The latter is, given the information we so far have, the better of the two. The speaker cannot expect to succeed at conveying (1) because his hearer will have no reason to suppose that there is just one tree, or one area in question, that is, that the speaker has one or several particular bunches of bananas in mind. Whereas, the speaker can sensibly be taken to convey that there are some ripe bananas.

We can distinguish between statements about particular objects, and ones about objects-of-a-kind, but this is not sufficient, of course, to establish the distinction between a subject-predicate rendering and an existential one. We have in addition to (1) and (2), (3) "there is one bunch of ripe bananas" (and by implication, the speaker says he knows where they are). Existential statements can be about particular objects, and unless certain rather special conditions obtain, we shall be unable to classify utterings as subject-predicate or existential once we have established that they are about particular objects. There are two considerations relevant
to the classification: whether the picking-out is a picking-out of
the object as of a kind, and whether the hearer knows already of
the existence of the object in question. Presentations and point­
ings do not pick out objects as of a kind, and for that reason take
the subject-predicate rendering. This is one case. The other comes
about when the speaker makes an utterance about a particular object
which, he believes, his hearer already knows to exist. He cannot,
therefore, intend to induce the belief that there exists such-and-
such an object, so his utterance cannot take an existential render­
ing. In all other cases, I think that we have no reason to adopt
either the subject-predicate or the existential as the one correct
rendering of the meaning of the utterance.

Consider this story about two hunters, speaker and hearer.
They have gone hunting because a bear killed one of their sheep on
the previous night. The speaker, we suppose, is the expert tracker,
while the hearer is able with bow and arrow. They intend to kill
the particular bear which killed the sheep. The tracker has followed
the tracks from the sheep-fold, and as they come in sight of the
bear, he points: "That's the one we're after". An existential con­
strual will not do. In expecting to succeed at communicating, the
speaker presumes that his comrade already has considerable know­
ledge of the object being talked about; it is mutual knowledge be­
tween the hunters that they are searching for a particular bear, and
the speaker depends on this knowledge for success in communicating.
My story exemplifies the epistemic situation characteristic of referring, but referring doesn't take place, because there is no functionally distinguished referring part. The relevant epistemic material is present, though, so referring can take place whenever some proper part of a speaker's utterance has the function of picking out the object of the belief the speaker means to convey. One may feel inclined to say that the bear hunter has referred: given the appropriateness of his intentions and expectations, it may not seem to matter much that his utterance has no parts. We should resist the temptation, though. A notion of referring which makes it a semantic function of parts of utterances is useful, but assimilating referring to making utterances which are in a broad sense 'about' particular objects will do no good at all; we already have 'statement about a particular object', and we should anyway have to introduce the subject-predicate/existential assertion distinction somehow.

We have the relevant epistemic materials, now, so we shall have a primitive case of referring and predicating if we can see how a speaker might produce an utterance, distinguishing functionally and intentionally a referring part and a predicating part. I will use the two characters of the earlier scenario, who are again hunting for a particular animal. This time they are hunting for two wolves which have killed one of their sheep. One of the wolves has a deformed paw, and also is very large and heavy, as wolves go,
and so leaves deeper and bigger tracks. The hunter expert at tracking has of course noticed this and pointed it out to his comrade. The hunters come on the wolves they are looking for, and the tracker points to one wolf (a large, old male wolf) and then distorts his hand to indicate the crippled foot. The force of this combination is: "That wolf has a crippled foot."

I shall suppose that the wolf-hunting episode follows fairly soon after the bear hunting. The speaker therefore can count on his hearer's remembering how the pointing gesture was used on that previous occasion, and while it is sufficiently obvious that it is being used differently this time, something will carry over. The pointing gesture as a whole utterance was used to mean "That one is the one we're after." One way of showing, in our paraphrase of what the speaker meant, how the carry-over is helpful is thus: "That one [is the one we're after and it] has a crippled foot." The hearer may consider some rendering like this, and of course should reject it since they're after both wolves, not just the one. Part of the meaning brought from the earlier occasion will be ignored as irrelevant, leaving "that one ..."). The hand-gesture intended to suggest a crippled foot adds on pretty easily once the first part is understood: "That one has a crippled foot."

Our speaker has used the pointing gesture he had used previously, thus giving the hearer some additional information about speakers' intentions. By using the gesture in a way naturally
connected with its previous use, the speaker gains an advantage because he may suppose that his hearer takes the second use to be like the first. While, should he use the gesture in a quite different way, he would probably mislead his hearer, and in any case, would lose the help of a precedent. But, using an expression which has a history is not necessary to first cases of utterings with structure. An uttering could have two parts distinguished functionally neither of which were used in the language-group before. The range of things which are mutual knowledge with speaker and hearer will have to be enlarged in other ways, though, if there is no dependence on past utterances.

Our speaker intends something like "That one has a crippled foot," by his utterance, and in the circumstances set out, it seems reasonable to suppose he will succeed at conveying that. The function of the pointing gesture is to single out some one object, the function of the distorted-hand gesture is to say something about the object picked out, and the speaker intends them to function that way. The parts are distinguished naturally as two gestures and functionally as something like referrings and predicatings; together they have the function of getting the hearer to believe that the wolf has a crippled foot. If I can show precisely how the bit of theory about parts of utterances developed in Section 2 comes to bear, this case will be a logically primitive instance of an utterance's having internal structure.
A part P of some meaningful whole utterance U is meaningful for a speaker S if S, via the Gricean intention, intends that P have a certain role in realizing his U-intention. When one specifies what sort of role P is intended to have, one specifies what sort of part it is. I suggest these as the referring and predicating intentions. Referring: by some whole utterance U, S intends his hearer to come to believe that some object X is \( \emptyset \); he refers to X with a part P of U when he intends in the Gricean way that P make known to his hearer that X is the object of the belief conveyed by U as a whole. P is meaningful, and has the grammatical function of a subject-term, and what P refers to connects with what it means. Predicating will work in a suitably related fashion: in uttering some other part P', the speaker intends in the Gricean way that his hearer come to believe that the object of the belief conveyed by U and picked out by P has the property \( \emptyset \).

Notice that referring and predicating are explained in terms of the contribution a term makes to the whole utterance in which it is incorporated -- one cannot refer or predicate except in the context of a whole utterance. Austin\(^6\) expresses this essentially Fregeian point by calling reference and predication ancillary speech acts. The logical relations between the referring and predicating terms are a special case of the logical relations which hold between any set

---

of two or more terms whose semantic functions allow them to form whole utterances when combined. The problem is a familiar one: how do the parts of whole utterances combine to make a whole utterance? The theory presented here suggests that terms or parts are incomplete, and whole utterances complete, in the sense that the description of the semantic function of a part P makes essential reference to the containing whole utterance U, and, implicitly, to the remainder, U-P, which is presumed to serve some complementary semantic function. For example, in the context of a subject-predicate utterance, a full description of the referring part refers to the predicating part, and to the whole utterance; the description of the predicating part is parallel, containing references to the whole utterance and to the complementary referring part. Very weak structures, e.g. two whole utterances joined by the conjunction 'and', needn't have internal logical relations to each other. That utterances are joinable in this way needn't appear in a definition of 'whole utterance', but can be explained entirely by the meaning of the word 'and', or more precisely, by the rules for joining whole utterances with it.

The wolf hunter who pointed and then made a distorted-hand gesture has produced a meaningful whole utterance with meaningful parts. The two parts are distinguished physically by their forms and intentionally by the speaker's intentions for their function; they refer and predicate respectively, and insofar as they do so,
have a suitable logical interdependence; and the speaker could use
them again, if he chose to, in other utterances to mean more or less
the same as in the first. He could, in other circumstances, and
with different hearers, use the same set of acts to communicate the
same thing, or he could use the pointing gesture independently of
the distorted-hand gesture, and vice versa, coupled with some other
suitable act which refers or predicates. I conclude that this is
a logically primitive instance of an utterance's being a referring
and a predicating, and, a fortiori, a logically primitive instance
of an utterance's having internal structure.

and Grammar

I shall apply Bennett's revision of Lewis' analysis of conven-
tion to generate a logically first instance of an utterance's hav-
ing a conventional meaning, conventional structure, and convention-
ally meaningful parts.

A speaker and a hearer have a co-ordination problem since a
speaker wants to be understood and a hearer wants to understand
what the speaker says. The speaker may intend by his utterance
simply to communicate a certain belief, or he may intend to communi-
cate a belief in a certain sort of way, with subsidiary intentions
for elements in his utterance. Upon application of Lewis' analysis
of convention, the first kind of case will give conventional mean-
ings for whole utterances, and the second, depending on detail, will
give grammatical conventions and word meanings. The first is straightforward: speaker and hearer have a problem because the speaker wants to communicate that p and the hearer wants to understand. Suppose situations arise fairly often in which a member of the group wants to communicate that p, and so some utterance U becomes regularly used for that purpose. When the members of the group begin to use U to mean that p because it is mutual knowledge among them that that is what U has meant in the past, U's conventional meaning is p. The second is more complicated.

I shall proceed by showing how examples of the genus structural convention might develop, marking the species grammatical convention and part-meaning convention by further differentiation in function among the parts of utterances.

Let us stay with the pointing gesture, but vary the gesture or noise which has the predicating function. The meaning of the whole utterance will therefore vary only with the meaning of the predicate-part and with what object the pointing is used to pick out. What will remain constant is the function of the pointing part -- referring -- and the distinction in function between the pointing and the predicating part. We can suppose that this utterance-type becomes established as a convention among the members of the language-group. They have a certain sort of recurring communication-problem: to pick out particular objects and to say something about them. If oneself and one's hearer are in the presence of the object,
one points to it and predicates something of it using some convenient and suggestive action as the predicate-part. Pointing has acquired a conventional meaning — in fact, it has here much the same meaning as it has for us — and pointing plus a predicate part is an utterance-type which has a conventionally assigned use.

This utterance-type has a subject-predicate structure of a sort, but structure and meaning haven't been clearly separated yet. Let us take as defining of a purely grammatical convention that it specifies a conventionally recognized relation between ordered sets of words or meaningful terms each of which belong to some appropriate grammatical category. The convention doesn't mention which words happen to be used, but specifies a relation between any words of the appropriate categories. So before we can formulate a purely grammatical convention, we must have means to group words into grammatical categories. Our example convention isn't a manifestation of a grammatical convention, therefore, but of some simpler structural one not handily divisible into meaning and grammatical components. The pointing gesture has a referring function, but since it is the only referring term the convention governs, the convention should not be described in terms of the role of subject-expressions and predicate-expressions, but in terms of the role of pointing gestures and the expressions which follow them. However, once we have a range of expressions all of which can function as subject-terms, and a range of expressions for predicate-parts (though not all combinations
need be allowed), what all the whole utterances so formed will have in common is grammatical form. Choosing a simple form, let us say that the first term in a two-part utterance is the subject-part and the second the predicate-part. Identity-statements might be distinguished by a different form to reduce ambiguity.

Probably the easiest way to get more subject-expressions is to introduce names. Descriptions could be introduced first, I think, but it would be hard work to construct cases sufficiently resistant to an existential rendering. The use of names to pick out for one another mutual acquaintances will be, just like the pointing gesture, a case where an existential rendering is inappropriate. Suppose one of the persons in the language-group has the habit, while engaged in any kind of sedentary work requiring attention, of involuntarily making the sound 'mm-mm'. Someone might use this habit of his, and this sound, to refer to him. The hearer is rightly impressed with the utility of doing so, and uses it himself on occasion. Someone else may have some characteristic trait which lends itself to being used as a referring term, or which makes some term appropriate and suggestive; many people will soon have names.

The group now has a grammatical convention, one which we might describe as follows. The combination of an expression of the subject-term type followed by some predicate-term type is a conventionally recognized utterance-type; a subject-term has the function of picking out an object, and a predicate-term the function of saying
something about the object picked out.

I suspect that a grammatical convention for subject-predicate utterances could develop without the accompanying development of conventionally meaningful terms. Using a term like the pointing gesture as a common element around which a convention begins is certainly the easiest way to proceed, but it seems possible that a sufficient range of subject-expressions could be available, and be used in such a way that a simple, linear structural convention could develop without any part acquiring a conventional meaning. It is possible, that is to say, for the categories of subject-expression and predicate-expression to have use in describing a language without there being terms which are recognized as conventional instances of those categories. In order for this to happen, though, the group of communicators would need powers of imagination sufficient to create new terms when necessary, in numbers great enough to avoid the repetition which would naturally lead to a convention. And they would need abilities to see connections between utterances along a dimension which seems for us not so easy to see, a dimension of purely functional similarity and difference. So I think simple grammatical conventions could develop without conventionally meaningful parts. I don't know if words can develop without conventional grammar, not because something paralleling the previous case can't happen — it can — but because, as we shall see, it seems odd to call what results a part with a meaning. It will be as if there
were a certain kind of systematic ambiguity naturally arising. Suppose that the sound 'mm-mm' became associated with the person in the previous story, in such a way that it was sometimes used in the Gricean way as a term which refers to him, and sometimes as a predicate indicating the property of being like him. Could we then say that 'mm-mm' was a kind-word, used to refer to or talk about mm-mm- mers, but without its being embedded in any structural convention? If we take 'mm-mm' as a kind-word in this way, we just will have described a structural convention, since 'mm-mm' is used as either a subject-term or a predicate-term, in utterances which either refer to mm-mm, or attribute the property of being mm-like. If we are to avoid this, the term must be used in a large enough variety of Griceanly effected grammatical structures to justify us in thinking that no convention is in effect. That is, it must be used in ways which, if it should be possible to bring them under a convention, would result in one too complex to suppose the speakers and hearers actually to be using it. While this may be possible, the demands on the imagination and intellect of our communication-group will be so severe as to make it highly unlikely that they have any terms like 'mm-mm'. Further, what can we say about the meaning of 'mm-mm'? Not much. Because it can (must) be used in such a great variety of ways, it can't be a noun-phrase, or a verb, or any other likely sort; otherwise there will be enough uniformity to cause us to suspect a convention. The meaning of the term must be very
diffuse, since its meaning must be unrelated to the grammatical functions it performs. Its 'meaning' would probably be unlearnably complex.

The convention I've outlined for subject-predicate utterances makes a pointing a subject noun-phrase, whereas we should like it to be classified more generally as a noun-phrase. We have not yet distinguished the utterance-relative category subject noun-phrase and the non-relative category noun-phrase, but we need the distinction. We don't want to be forced to say that the meaning of a term differs depending just on whether it is used to refer or to predicate. Yet, unless our convention will let us classify a term as a noun-phrase, rather than more restrictively as a subject noun-phrase, we will be forced to treat its occurrence in the predicate as a new word. The problem could, I think, arise and be treated before we consider word-meaning conventions to obtain, but it arises most pressingly when there are word-meaning conventions to be described. For instance, suppose the men of our tribal group belong to a bear totem. One can say (as an injunction, directed at a particular person) "Bear-man, be brave", as well as being able to say, "He is a bear-man."

Now, unless our group has some sort of speech activity which demonstrates that they recognize this difference, we must credit them only with the less general convention -- we shouldn't attribute more complex conventions than necessary to explain their
behavior. The sort of evidence which shows that they do have the more general convention is precisely that they do use some terms in other than subject position, while preserving continuity of meaning. And when such terms become part of the repertoire of the group, they have begun to advance to a more sophisticated way of grouping words into non-relational grammatical categories, grouping them only secondarily, in the context of particular utterances, into the relational categories of subject-term and predicate-term.

The typical benefits of the interaction of grammar and words can now appear in full. Once we have some grammar, new predicate-parts with conventional meanings develop more easily. There is a conventional means for marking a part as having a descriptive, a predicative function — it appears in the predicate-place in an utterance with conventional subject-predicate structure — so one can introduce new terms ostensively and most of the problems traditionally posed, e.g. how the hearer can be expected to understand what the speaker means, are removed because we have a grammatical convention which assigns functions. If, when queried about N, I say 'N hunts' and, since I am introducing a new term 'hunts' I mimic someone hunting, I have a good chance of being understood because my hearer knows how the term ought to function. New terms will come fairly easily now, and we can suppose that our group develops a set of descriptive terms as large as they find useful.

The group has already a very powerful language, and new
conventions develop easily. It is a major virtue of Lewis' analysis that it shows how speakers' intentions can lead to new linguistic conventions and can alter and extend established ones.
Chapter 3
A Gricean Theory of Reference

1. Intending to Refer

I shall argue that referring should be analyzed in terms of speakers' intentions rather than in terms of the grammatical structure of utterances or the meanings of their parts. At least one consequence of such an analysis is immediately attractive: by invoking speakers' intentions, we shall be able to explain the acceptability of a large class of utterances which have resisted explanation but which obviously are acceptable. This class contains, for instance, all utterances in which the term or terms which purport to refer fail to denote, yet where it is clear the speaker is talking about a particular object. One way, and it seems the only way, in which non-denoting referring terms may function to pick out objects is through a speaker's intending them to. This intention, I shall argue, is of the form Grice uses in his analysis of meaning, viz: a speaker in uttering U intends to affect his hearer in some way A, intending that this come about at least in part through the hearer's recognition of the speaker's intention. I shall begin by showing the sufficiency of a certain set of speaker's intentions in

1. A definite description denotes when it is true of just one object, and in a slightly extended sense, a name denotes when it names, as a matter of convention, just one object.

effecting reference; that set is also necessary, though I shall
defer until later argument for this latter claim.

Some referring terms fail to denote because they are true of
more than one object and some fail because they name more than one
person. Such failures occur rather often. In ordinary conversa-
tion we neglect to ensure denotation just because. I conjecture,
we are sure that our hearer will nevertheless know which object we
mean. For instance, while confronted with a group of several dogs,
one may say to a companion, "The brown dog is ...". Now of course
the brown dog one means to pick out is the one which is a member of
that group. That isn't what has been said, but one may reasonably
expect the hearer to understand which dog is meant. Hearers do very
often know which object is being talked about even though it has not
been specified both explicitly and uniquely; this is simply a fact
to be adduced. To explain it, though, we must suppose that a hearer
infers the identity of the referent from what the speaker says and
from context. The hearer must infer which object it is the speaker
meant to pick out; he cannot tell which dog is being talked about
just by understanding the referring expression 'the brown dog' since
it does not denote. Similarly, while we usually refer to people
using their given names, such names are often ambiguous many times
over; our use of proper names would be utterly mysterious if spea-
kers' intentions weren't present to remove ambiguity. I offer as
a final example of referring terms which fail to denote those which
have no conventional referents whatever. We have a few such expressions, e.g., 'the whatsit' and 'the whatchamacallit'. They are conventionally used to refer to whatever object the speaker desires, which is to say the object he intends to pick out.

In the cases so far discussed, we may secure denotation for definite descriptions by adding appropriate restrictive clauses, e.g. "The brown dog that is in front of us is ..." Names are somewhat different, but when we say, "John Smith of the University of British Columbia is ...", the modifying phrase performs the same function as a restrictive clause, and so it is part of the referring term. However, not all problematic instances arise through incomplete specification of the referent. It is sometimes false that the object the speaker means to pick out is identical with the object the expression he uses denotes. A speaker may, for instance, be mistaken as to which object satisfies a certain denoting expression. If he refers with that expression, the object he means to pick out will not be the object denoted and he thus unknowingly violates convention. One may also knowingly refer to some object using an expression whose conventional referent is a different object; here one flouts convention. I shall discuss these possibilities in order.

When I believe mistakenly that 'the Ø' denotes X I may, should the occasion arise, use it to refer to X. You as the hearer will try to determine which object I mean to talk about, at least when it is clear that I can't be talking about the Ø. Sometimes this is
very obvious: I might say, "Your eldest son certainly has bright red hair," when it is really your youngest son whose hair is so noticeable. You would probably conclude that I meant to talk about the youngest son, and that I believe falsely that he is the eldest. Of course it will not always be evident when a speaker believes falsely of some object that it instantiates the denoting expression he uses; when it is not, a hearer will suppose, wrongly, that the denotatum is the object of reference.

I think that one can knowingly use a referring expression deviantly, flouting convention, just when one expects to be understood despite one's deviant performance. The truth of this conjecture follows from a Gricean theory of reference, but since I haven't yet argued for such a theory, I shall merely point out that we sometimes do use expressions deviantly, confidently and often correctly expecting to be understood. Poets do this regularly in metaphor -- a considered flouting of meaning-conventions is one of their tools. Examples of metaphor are easy to find, but I need an example of an expression used both metaphorically and referentially. As it happens, I enjoy fishing and so I occasionally read the magazines catering to fishermen. They often contain articles by non-professionals, some of them given to a rather purple prose, as for instance this: "At last! The silvery jewel was in my net!" 'The silvery jewel' refers to a particular fish, but a fish is not a jewel though it may be silvery. The writer is using the phrase
deviantly, yet because of context no one will have difficulty understanding what he means.

My examples show (1) that reference sometimes occurs even though the referent is incompletely or incorrectly specified, (2) that in such cases we have evidence that the speaker intended to pick out a particular object for the hearer, and (3) that the hearer, if he succeeds in determining the referent, does so by inferring that some object X is the object the speaker meant to pick out. There is a certain conceptual fact about intending which, together with (1) through (3), will let me conclude that reference is effected in these cases in a Gricean way. If one intends to do A, one expects to succeed at doing A. Applying this truth about intending to (2), we find that if a speaker intends to pick out a particular object for his hearer, he expects to realize that intention. In circumstances such as those my examples illustrate, though, there seems to be just one way in which a speaker may expect his intention to be realized, namely by the hearer's recognition of that intention. Now if one knows an action to be composed of two parts A and B, neither of which may occur without the other, one may not intend to perform A without B nor B without A. In order to apply this fact to the case at hand, I need a slightly extended sense for 'action', one which applies to anything one may intend to do. Then, since in my examples a speaker cannot pick out an object for a hearer unless the hearer recognizes the speaker's intention to do
so, it follows that a speaker cannot intend to pick out an object for his hearer without also intending the hearer to recognize that intention. I conclude that in the cases I've been discussing, reference occurs because the speaker intends to pick out a particular object for his hearer, in part by means of the hearer's recognition of that intention, which is to say that reference is sometimes effected in a Gricean way.

It is an 'intend'-to-'refer' conditional which has, indirectly, been the subject of my discussion. Sometimes, we have seen, reference is effected just by the speaker's using some part of an utterance with the intention to refer. In order to conclude from this that any time a speaker has the proper intentions in speaking, he refers, we need only to establish, which in fact I think I have, that in the cases brought as examples, it was the speaker's intention in speaking which effected reference. If it was, then reference may be effected in that way, and some form of an 'intend'-to-'refer' conditional is true. Now, since an analysis is composed of a pair of conditionals, I shall need a 'refer'-to-'intend' conditional in addition to an 'intend'-to-'refer' conditional. In Section 2 I develop a preferred form for the referring intention, as I shall call the antecedent in the 'intend'-to-'refer' conditional, and defend it against objections. That reference is sometimes effected in the Gricean way gives us no reason to suppose it is always effected in that way, and I shall need different and
stronger arguments to establish the truth of the 'refer'-to-'intend' conditional, or in other words to show that all referring is effected by Gricean intentions. I do this in Section 3. In Sections 4 and 5 I consider the epistemic aspects of referring, that is to say, what a speaker must know about the world if he is to be able to refer to some particular object.

2. The 'Intend'-to-'Refer' Conditional

I have been describing a speaker's intention to refer as his intending to pick out an object for a hearer. That form is conveniently descriptive, but it clearly won't support any theoretical burden: 'picking out for a hearer' is itself in need of explanation. With this in mind, I offer the following as the penultimate version.

The referring intention:

In producing a meaningful whole utterance U with a proper part D, a speaker S intends his utterance of D to bring his hearer H to believe that X is the object of which S predicates some property with the remainder of U, intending this to come about in part by H's recognition of S's intention.

(Suitable alterations will accommodate the non-indicative moods.)

When a speaker realizes his referring intention, he performs a fully successful act of reference. He may fall short of complete success in either of two ways, however, only one of which engenders a failure to refer. The referring intention has an audience-
directed and an object-directed component; and while success or failure along the first dimension concerns hearers' uptake, failure to refer occurs only through a failure along the second dimension. A speaker fails to refer when he fails to pick out an object, but before we can explain how failures occur, we need to explain what is right about successes; that is, we need an explanation of the epistemic connection between referrer and referent. The presence of such a connection explains how a speaker has an intention 'directed at' an object, and the failure of the connection explains how a speaker fails to refer. For various reasons I shall defer this matter to Section 4.

An 'intend'-to-'refer' conditional tells us that when the antecedent is true, reference occurs, so both the referring intention and its success-conditions need to be part of the antecedent. The conditional below displays the correct form, but I do not treat success-conditions until Section 4.

If a speaker S, as a part of producing a meaningful whole utterance U, utters D with the referring intention, and if he succeeds thereby in picking out the object of his intention, he refers to that object with D.

There are two criteria of adequacy for a referring intention in a Gricean theory: it must meet both (1) the theoretical demands
placed on any analysis of what a referring term does, and (2) the additional demand for detectability arising from the hearer-directed aspect of Gricean intentions. 'Picking out for a hearer' was suggestive of the effect the speaker intends to achieve, but as an explicit rendering of that effect, I've offered 'the belief the speaker intends to induce in his hearer', i.e. that X is the object which the remainder of U (U -D) is 'about'. We apparently have no term for the genus semantic functions which, coupled with a referring, yield a whole utterance, so perhaps the best we can do is give a disjunctive listing of the functions we recognize, e.g. for assertions attributing a property to, and for imperatives requesting that something be done to, the object of reference. We need reference to U and to the function performed by the remainder of U because a term cannot refer in isolation -- it must be coupled with another term which serves some complementary function, the two together making up a whole utterance. A 'picking out' of an object is a referring just when a speaker 'picks out' with a certain purpose, namely to go on to do or say something about that object in the course of producing a whole utterance. The speaker intends his hearer to believe that X is such an object with respect to U.

Should a speaker have the referring intention in uttering some term D, a hearer will almost always have evidence letting him detect that intention. His best clue that referring is taking place lies in the grammatical form of U. If it is one which conventionally
signals that referring is taking place, he may take this as good evidence that the speaker meant to refer; his remaining problem is determining which the referent is. The hearer's grammatical knowledge solves another problem also, since we may suppose he has a working knowledge of the predicative function, or of its counterpart in imperative utterances. But, grammatical convention may be flouted, so in deviant cases a hearer may not depend on the explicit grammatical form of the utterance he tries to understand. I don't think there is anything systematic about the sources of hearers' information in grammatically deviant cases. We should have to examine each case to see whether a hearer has enough information to detect speakers' intentions; if he doesn't, the speaker should not have attempted to refer deviantly.

While we depend on grammar very heavily indeed to provide evidence of speakers' intentions, I think that referring is not essentially connected with grammar. When grammatical convention is flouted, referring terms are distinguished just by being uttered with the referring intention, though there is arguably some dependence on convention even in deviant cases. A Gricean theory of utterance-meaning and part-meaning might be capable of showing how referring and predicating can occur entirely without benefit of convention, and while I believe such a theory can be developed, my present aims in no way depend on it.

I know of only one objection to the 'intend'-to-'refer' condi-
tional; it was first raised, I believe, by Strawson as an objection to the sufficiency of Grice's analysis of meaning. My 'intend'-to-'refer' conditional is isomorphic to the 'intend'-to-'mean' conditional, so Strawson's counter-example bears equally on both.

Grice suggests this analysis of meaning:

A speaker S meant that p by or in uttering U if and only if he intended that his utterance of U be at least partially responsible for bringing his hearer H to believe (or think) that p, and this by means of H's recognition of S's intention in uttering U.

The 'intend'-to-'mean' conditional in Grice's original formulation allows a speaker S to have a certain perverse sort of intention while intending his utterance U in the Gricean way. Strawson describes such a situation in the following way. S intends to bring H to believe that p, so S creates a state of affairs C which is apparently evidence that p, in such a way that H will be sure to notice it. S does this knowing that H is observing him, and knowing also that H does not know that S knows he is being observed. S knows that H will not therefore take C as genuine evidence that p, but expects that H will take S's creating of C as a reason for


thinking that S intends to induce in him the belief that p. S does intend to induce that belief, and H recognizes that intention, as S intends him to. But this is not all. H believes that S did not really intend H to recognize S's intentions; H thinks that S was trying to trick him. S knows this; he intends that H believe that S did not utter U (create C) with the Gricean intention. If we suppose H to believe that S was really trying to help him, H may come to believe that p because of his trust in S and his recognition of S's intentions. The conditions have been satisfied, but clearly this is not a case of meaning. In general, Grice's analysis does not bar cases with this structure: the speaker S has the Gricean intention i, but also has some other intention i'. If i' is or contains an intention on S's part that H should think that S did not utter U with i, but that H should think that S wanted H to think that S did, we must refuse to allow that S meant something by or in uttering U.

Grice⁵ has attempted to specify the kind of intention which causes the problem, and then bars it by stipulation. Taking a different approach, Schiffer⁶ sees this sort of case as one in which the speaker has certain 'hidden' intentions, and suggests that the proper remedy is 'to get everything out into the open'. He offers


⁶. Steven Schiffer, Meaning, Oxford University Press, 1972, see especially pp. 30 - 42.
a variant on the original Gricean analysis which works to secure the result that S cannot utter U with the Gricean meaning-intentions while intending to delude his hearer in this special way. Schiffer suggests that S meant something by uttering U if and only if S utters U with the Gricean intentions and intends by his utterance of U to create a state of affairs E which is sufficient for S and H to mutually know that (1) E obtains, (2) that S intended that E obtain, and (3) that E is good or conclusive evidence that S did utter U with the Gricean intentions plus his intention to create E. The state of affairs E is S's utterance in context, which is anyway the source of H's information about S's intentions, so if E is rich enough to yield (2) and (3) we shall have secured our result. It seems clear that E very often does yield that result.

Schiffer introduces mutual knowledge as a technical term to describe this sort of situation: S knows that p, H knows that p, S knows that H knows that p, H knows that S knows that p, S knows that H knows that S knows that p, and so on. If it is mutual knowledge between S and H that E obtains, and if E is good or very good evidence that S uttered U with the Gricean intentions, we have ensured that S has no 'hidden' intentions of a sort to disrupt the functioning of the normal Gricean ones.

When Strawson first discussed his counter-example, he noted that the natural way to bar it would generate an infinite chain of intentions. What he had in mind was this. We may try to bar the
'hidden' intention by adding a third level of intention: S's primary intention is to do A, his secondary intention is that his primary intention be recognized, and his tertiary intention is that his secondary intention be recognized. But a Strawson-type counterexample can be generated at this level too, and so be barred by a fourth level of intention, be posed again, be barred by a fifth level, and so on. There is a certain similarity between the infinite chain of 'know that' statements to which 'mutually knowing that' is equivalent, and the infinite chain of intentions Strawson rightly takes to be unacceptable. But, Schiffer argues, certain sorts of mutual knowledge and so certain infinite chains of 'know that' statements are harmless. When under normal conditions two people sit at a table looking at one another, their doing so is sufficient to guarantee the truth of the claim that they mutually know that they do sit across the table from one another. Just because of their knowledge of the situation, they are each justified in claiming that he knows that the other knows that ... etc. to whatever level of complexity one wishes. The entire chain is generated by the facts of the initial conditions, so it is harmless.

I suggest that Schiffer's solution to the Strawson-type counterexamples is to be preferred both for meaning and for reference. It incorporates a device which ensures that just the relevant intentions operate, while achieving this end not simply by barring certain cases, but by explaining what it is about proper cases of Gricean-
type meaning which secures their propriety; it does this, furthermore, in a way which shows how a hearer may hope to detect improprieties. Incorporating Schiffer's analysis makes the referring intention rather cumbersome, however, so as a matter of convenience I shall continue using the older, simpler form.

3. The 'Refer'-to-'Intend' Conditional

Difficulties for the 'refer'-to-'intend' conditional arise both from the fact that one uses a part of an utterance to refer, and from the fact that in sharing a form with the 'mean'-to-'intend' conditional, it also shares some difficulties.

Parts of utterances raise special problems because, while Grice's analysis of meaning arguably applies to all whole utterances, it very obviously does not apply directly to parts. That is to say, we cannot expect successfully to analyse the meaning of any particular meaningful part of an utterance in the Gricean way. Should I assert "This fish is barely edible," do I have an intention in uttering 'barely' which is distinct from my intention for the whole utterance and which effects the contribution 'barely' makes? Surely not -- one depends on the word's conventional meaning and grammatically assigned function. Even a single-leveled and non-Gricean intention would have to contain a specification of 'barely's grammatical role, but I may not know what it is. One can construct grammatical sentences and in that sense know grammar, without being able to describe fully the grammar of the sentences one composes.
So, in order to show that if one refers one has the appropriate Gricean intentions, I shall have to show that referring is a special case, and that referring terms are unlike other meaningful parts just in that they always do achieve their effects in a Gricean way. I shall argue to this conclusion in two steps corresponding to the two parts of the Gricean intention.

If a speaker distinguishes some part of his utterance functionally as a referring term, he is able to respond appropriately and, usually, informatively to the query "Which did you mean?" Suppose he asserts, "The $\phi$ is blue." If we, as hearers, don't know which object the $\phi$ is, we may ask "Which is the one you say is blue?" or "Which is the $\phi$?" In asking this question, we indicate that we suppose the function of 'the $\phi$' to be that of picking out a particular object, and not, say, attributing the property of blueness. If we don't suppose that 'the $\phi$' has that function, we won't ask the question, and if we suppose of no part of the utterance that it has a referring function, we can express our lack of understanding only by asking, "What did you mean?" Similarly, if the speaker understands our question, "Which one did you mean?" he indicates that he recognizes that 'the $\phi$' functions to pick out an object, while if he doesn't understand, we may conclude that he doesn't distinguish 'the $\phi$' functionally as a term which refers. But a speaker cannot, I shall argue, be ignorant of which are the referring terms occurring in his own utterances.
Consider some French children playing at cowboys and indians. One says to another, "Stickemup", having learned this utterance from watching American western movies. It is a meaningful whole utterance, but because the child could not use 'up', 'stick' or 'em' in combination with other meaningful parts, meaning much the same by them in each use, they in fact are not meaningful parts of his utterance. As a consequence, in giving the order, "Stickemup" he does not refer with 'em' ('them'). Uttered by an English-speaking person, 'them' would of course refer to the hands of the person addressed, but the French child does not refer because, though he knows the fulfillment-conditions for his imperative utterance, it has no meaningful parts, and so there is no connection between its fulfillment-conditions and its structure. Whereas, to say that some part refers is to say something about the relation between structure and truth- or fulfillment-conditions. Appealing to the fact that the child's utterance is an English sentence, and in that sense has a conventional structure, to be correlated with its fulfillment conditions (which the child does know) will not do either. The child need not know that the sentence belongs to English, but suppose he does. Even this, I think, will not justify us in giving the child's utterance the structure which it would have were it uttered by a competent speaker of English. Neither the child nor his hearer know the meanings of the words which make up that English utterance-type, so the speaker does not depend on the
conventional meaning of "Stick 'em up" in English to implement his communication desires. The fact that his meaning something by his utterance does not depend at all on any knowledge he has of what the words in it mean shows that what he means has no relevant relation to what the utterance conventionally means: it is in the important sense an accident that the English sentence "Stick 'em up" means what the child means by his utterance of it. We should not, therefore, take facts about the meaning and grammar of the English sentence as determining in any way what the child means or does.

One may be ready to grant that the French child in my example does not refer, yet refuse to grant that referring may take place only if the speaker understands that the relevant term in his utterance functions referentially. A speaker might, so the objection goes, be able to produce and understand utterances with internal structure including terms which refer, without being able to identify them. He is able to use conventionally meaningful parts — words — in various utterances and in various combinations; his ability to do this is good evidence that he knows how to compose and understand grammatical sentences compounded from words whose meanings he knows. Such a speaker, the objector concludes, must surely be said to refer with the grammatically appropriate parts of his utterances, yet it does not follow that he knows which those parts are nor what their function is.
This supposition is false, I shall argue, since a speaker cannot understand new sentences he composes with the words he knows unless he also knows a certain minimum something about the connection between the structure of his utterances and their truth-conditions. What is it, after all, that distinguishes a speaker's knowing words and being able to use them in various utterances, from his having a repertoire of utterance-types which together specify a set of words? Surely, it is the ability to create and understand new utterances, ones he has not used or heard used before. If a speaker does not recognize the simpler aspects of the referring function of terms, he will be unable to create and understand new utterances whose internal structure is reference-and-predication. Suppose a speaker has the words 'the', 'cat', 'is', 'on', and 'mat' but has never heard or used them combined as 'the cat is on the mat'. Unless he knows that the two definite descriptions have referring functions, he will be unable to understand the sentence should someone use it, and he will be unable to utter it with assertive force only when he wishes to say of a particular cat that it is on a particular mat. We may conclude that any speaker who is able, using the words he knows, to create and understand new utterances whose functional structure is referring-and-predicating recognizes the referring terms which occur in those utterances.

Distinguishing referring and predicating is the first step in
what we might call the practical semantics of subject-predicate utterances: a speaker cannot produce utterances of that sort without distinguishing referring and predicating, but neither does he nor his hearer need to be able to make finer distinctions in structure and function. Someone who recognizes the referring terms and the two-place predicate 'is on' in 'the cat is on the mat' will not be critically lacking in practical knowledge of truth-conditions should he not know the functions of 'the' and 'cat' taken separately. Similarly, should our speaker produce the utterance 'the clouds move slowly', he may be said to know its truth-conditions as well as the meanings of the words, even though he may not know the grammatical role of 'slowly'. If queried, "Slowly?" he may be able to say only "It means they don't move rapidly," which is not to explain the word's grammatical role. By contrast, if our speaker were asked about his 'cat' utterance, "Which one do you mean?" or "What does 'the cat' do?" he would be able to respond appropriately.

I have been arguing that a speaker is always able to identify the referring terms in his utterances; he knows their function, at least to the extent required to understand and respond to the query, "Which one did you mean?" The question's appropriateness suggests that in addition to knowing that some term serves to pick out a particular object, the speaker intends that it shall. Suppose a speaker whose grasp of the communal grammar isn't completely sound produces the utterance, "It runs is tired," instead of the correct,
"The one who runs is tired." To a hearer who expresses lack of understanding, the speaker will indicate (with a pointing) that 'it runs' was to pick out that person, which is to say that the speaker intended to pick out a particular person, namely the one he pointed to. Someone who believes a term he uses has the function of picking out an object will be ready to make clear to a hearer what is picked out, thus indicating what function he understands the term to have. If a speaker is able to explain his performance in this way, we may conclude that he intended his performance to be a certain sort, and by application, we may say that when referring occurs, a speaker intends his utterance of some term to be a referring. There is a principle at work here which connects the degree of intentional articulation of an action with the distinctions the agent would make among the parts of his action for purposes of criticism. If a person is able to criticize or explain some part B of an action A he performs, he shows that he takes a proper performance of A to include a proper performance of B. If he intends to perform A properly, therefore, he intends to perform B properly also. But should he be unable to distinguish some part C in criticism, he cannot have an intention with a distinctly C-directed aspect, since he knows nothing about how he should perform C. This principle, applied to uttering-and-meaning as a special case of things we may intend to do correctly, yields my conclusion.

I wish to point out that this conclusion is relatively weak
since it shows only that a speaker intends that referring terms in his utterances shall function to pick out objects. Speakers' intentions are not particularly important unless we may correctly identify the object of reference with the object the speaker means to pick out. There are, however, certain cases of what we commonly call referring, but in which there is no object such that the speaker meant to pick it out. These are not really referrings, as I shall argue in Section 4; they differ radically in epistemologically fundamental ways from central cases of referring. In all remaining cases we may identify the referent with the object the speaker meant to pick out, but the question is, do we need to? In Section 1 I pointed to certain facts about our practice providing evidence that we do need to make the identification, namely that we suppose referring sometimes to occur even though (1) the referring expression fails to denote, or (2) there is a discoverable disparity between the denotation of the expression and the object the speaker meant to pick out. (2) is the problematic case because on no other occasion is the object denoted not identical with the object the speaker means to pick out. We usually do identify the referent with the object the speaker meant to pick out, and in doing so we gain in intelligibility and utility, sacrificing nothing except the doctrine that referents are determined purely by the meanings of referring expressions. The reasons brought in defence
of this doctrine are usually epistemological, but I shall argue in Sections 4 and 5 that epistemological pressures push us in quite another direction, namely to holding that what the speaker intends is primary.

Losses in intelligibility and utility arise from letting convention determine reference. We should have to say, for instance, that when some speaker uses a denoting expression 'the Ø' to pick out an object X which he mistakenly believes is the Ø, he refers to the Ø and not to X. What he says will almost certainly be false of the Ø, since he meant to be talking about X, and chose the remainder of his utterance accordingly. Our speaker will also be uttering a falsehood if we construe him to be referring to X, since he is misdescribing X as 'the Ø', but an utterance falsified in this way may often be informative and important. What prompted the speaker to communicate was his desire to tell us that X is P, usually presuming that we already know of the existence of X. He therefore may realize this aspect of his communication-intention, even though uttering a falsehood, and we gain whatever benefit there is in knowing that X is P.

Once the referent has been identified with the object the speaker means to pick out, the argument to the recognition-of-intention feature goes easily. If a hearer is to understand what

is said, he must determine which object it was the speaker meant to be talking about, and in so doing, the hearer recognizes the speaker's intention to pick out that object. For the speaker's part, he expects his hearer to recognize what he intends by some referring term. He knows that without hearers' recognition he cannot expect to pick out a particular object for his hearer, and therefore he cannot and does not intend to do the one without the other. I conclude that whenever a speaker refers, he intends to pick out a particular for his hearer in part by his hearer's recognition of that intention.

There are two objections against which my conclusion should be secured, the first arising as a threat that the Gricean intentions do too much, and the second that they do too little. The first is that my arguments may imply an obvious falsehood, namely that we effect the meanings and functions of all parts of our utterances with Gricean intentions. They haven't that consequence, however, and furthermore they show why it is that except in some special cases, speakers depend essentially on conventional meanings. A referring term connects to the world in an outstandingly obvious way, in comparison with some other sorts of meaningful terms. Even in those cases in which a hearer has no relevant information except the meaning of the referring term to guide him to the referent, he could have more information by knowing more about the speaker's present concerns, and he would alter his judgement of which the
referent is should new information become available. By contrast, what sort of evidence does a hearer have which might let him determine what 'the' in 'the cat' did? or what 'barely' in 'that fish is barely edible' does? He very seldom has any at all. A general knowledge of a speaker's concerns and situation in the world isn't enough to determine the meanings and functions of most terms; in very few contexts will there be the right kind of information. For instance, the circumstantial evidence a hearer has will not let him decide between construing 'barely' in the sample utterance as a simple negative, as meaning roughly 'unpleasant but', 'only somewhat', or 'more or less'. But the word does have a precise meaning when used by a competent speaker of English. Whenever information becomes available, however, speakers' intentions may be relevant; this happens in poetry, for instance, the enterprise itself providing a context in which a hearer has reason to watch for convention-flouting uses and evidence to bring to bear in detecting speakers' intentions. When speakers can expect meaning-intentions to be recognized, they may effect the meanings of their words in the Gricean way. This explains both why we can sometimes flout meaning-conventions, and why in the great majority of cases we cannot. I conclude that speakers almost never effect the meanings and functions of all parts of their utterances in the Gricean way; since they may almost never expect to realize such intentions, they do not usually have them.
Second, I need to answer those objections to Grice's 'mean'-to-'intend' conditional which also bear on referring. There are two, but the most serious of these for Grice's theory of meaning does not affect referring at all. Grice says that when I mean that p by an utterance I produce, I am attempting to induce the belief that p in my hearer. But we may quite obviously mean something by what we say without intending to produce a belief, e.g. in conjecturing. Grice has amended his account in various ways, though perhaps the best is substituting 'induce the thought that p' for 'induce the belief that p'. When I refer, though, I do intend to induce the belief that X is the object to which I attribute a property or order something done. If my hearer came to believe of the wrong object or of no object that it was the one I meant to pick out, he would thereby have failed to understand me, so I must intend to secure his belief.

The second objection to the necessity of Grice's analysis of meaning does bear on referring. When we specify a speaker's intention in the Gricean way, we make reference to a particular hearer, yet pretty obviously we make meaningful utterances addressed to no particular hearer, and sometimes in the entire absence of hearers. A speaker who presents a technical paper before an audience he knows contains laymen as well as persons with a relevant expertise may intend only that whoever has the proper background shall recognize his intentions, since he knows that only some members of the
audience can recognize them, and he doesn't know which they are.
Soliloquy is in one way similar, since there is no particular hearer which the speaker means to affect. In addition, though, we should want to deny that a speaker had meant something by an utterance if there could not be a hearer with an epistemic endowment enabling him to understand the speaker's intentions. I think that someone speaks meaningfully in soliloquy just if he correctly believes that there is or could be some person who, were he present, would understand the speaker's remarks and recognize his intentions. This solution for audienceless speakings applies directly to reference.

4. Referential/Attributive

It is characteristic of a Gricean theory of reference that the referent is identified as the object the speaker intends to pick out. Even in the absence of such a theory, though, it is clear that many of our uses of referring terms give convention a guiding but not a determining role: we often take the referent to be not the denotatum but the object the speaker meant to pick out. Donnellan \(^3\) calls this use of referring terms, or one very like it, the referential use. He thinks that definite descriptions have two distinct uses, the referential and the attributive. Used either way, a definite description serves to single out some one object and is an instance of what we would ordinarily call 'referring'. The

characteristic difference between the two, Donnellan suggests, is this: when an expression is used referentially, the truth of the utterance in which it occurs depends partly on what object the speaker meant to pick out, whereas when an expression is used attributively, the truth of the containing utterance depends on which object that expression denotes. Donnellan thinks and I agree that the attributive use is not really a referring, however, so I shall take 'referring use' and 'referential use' to be interchangeable, reserving 'attributive use' for that other sort.

I shall make important use of the referential/attributive distinction, since the epistemic demands on the speaker differ sharply depending whether he speaks referentially or attributively. These epistemic differences, furthermore, are reflected in the effect convention has on our freedom to use terms deviantly. The two uses differ in the way convention enters in determining or helping to determine which object is picked out — in speaking referentially we may use but are not bound by convention, while in speaking attributively, we cannot deviate from convention. My explanation of the distinction should not be taken as a report of what Donnellan meant; though I think my understanding of the referential use agrees well with his, I believe we differ on what makes a use attributive. In addition, I broaden the range of application of the distinction to include any term whose function is, as we would commonly say, referring. I shall say that such terms, so used, designate
objects; reference and attribution are exhaustive of the class.

There are two important differences between referring and attributing. (1) When an expression is used attributively, the conventional referent is necessarily the object designated, if there is one. This is not true of the referential use. And (2), if someone refers to an object, that is to say, if he uses an expression referentially, it follows that he has a relational belief about that object. Whereas, if he uses an expression attributively, we may conclude only that he has an attributive belief about the object picked out. For the remainder of this section, I shall be mostly concerned with the first feature; I investigate the second in Section 5.

When we identify the referent as the object the speaker meant to pick out, we characterize the use as referential. A speaker very often has some particular object he wants to talk about, and when he does, which expression he happens to use to pick it out is in a certain way inessential. Whenever, so far as a speaker's own purposes and epistemic constraints allow, he could just as well have used some referring expression other than the one he did use, he will be speaking referentially. A speaker may be constrained in two ways when he refers: (1) he may not wish to misdescribe the object of his reference, and (2) he may believe he will be unable to secure hearers' uptake unless he uses a conventionally correct referring expression. Constraint (1) does not bear on reference,
but on truth: should a speaker misdescribe the object he means to
pick out, he will utter a falsehood about the object he meant to
pick out. Constraint (2) bears not on reference but on hearer's
uptake. In neither case is the speaker constrained to be referring
to whatever is the denotatum of the expression he uses. The con­
ventional reference of the expression is inessential in determining
the referent, and in referring a speaker is not bound by convention
except circumstantially.

When one uses a designating expression attributively, the
object designated is necessarily picked out via the meaning of the
expression one uses. Only one kind of convention-flouting may occur
in attributive uses therefore, namely flouting of meaning-conven­
tions. This sort of deviance is much more demanding of hearers'
collateral information than deviant reference, and it is corres­
pondingly less common. Though I shall not in fact make the sub­
stitution, generality would be achieved by using 'the conventional
referent of one's designating expression construed as meaning what
one intends by it' in place of 'the conventional referent of one's
designating expression'.

A simple example of the attributive use involves a situation
in which the speaker's only access to the object he designates is
through the relation of conventional reference. Suppose that some­
one X believes that there are spies, and that they are 'a threat to
national security'. If he is teased about his belief, asked if even
the shortest or fattest spy is a danger, he may assert "Even the fattest spy is a danger." There is no particular person whom X believes to be the fattest spy; X simply has good reason for believing that 'the fattest spy' is uniquely instantiated. X can in no sense have 'the fattest spy' in mind, since having a person in mind requires at least that one know a number of the important things about him; when this is true, there is some possibility of divergence between the object designated and the denotation of the expression used. There are two sorts of designating expression which fit this pattern, ones formed with superlatives, and ones which do not now but will in the future come to denote; each sort gives rise to cases in which we may be justified in believing that a definite description does or will denote, without believing of anything in the world that it is or will be denoted by that description. Thus, 'the first child born in the 21st century' and 'the Prime Minister of Canada in 1980' will at the present have only attributive uses: the object designated cannot be other than that the expression will come to denote.

One can, however, use a designating expression attributively while having a particular person in mind whom one believes instantiates it; unlike referring, though, one will not be intending to pick out that particular object or any other. One of Donnellan's examples shows this clearly. Suppose someone has murdered Smith in a particularly gruesome manner, and, since Smith was a very kind
and likeable person, we conclude that his murderer was insane. We say, therefore, "Smith's murderer is insane," using the designating expression attributively. Suppose too that we believe of some particular person, one whom we have clearly in mind, that he is Smith's murderer. Can we have been speaking attributively then? We can, because should we be wrong, and the person we have in mind turn out not to be Smith's murderer, we would not have uttered a falsehood. The truth of our utterance is assessed taking Smith's murderer, whoever he actually is, as the designatum. Had we been speaking referentially, we would have said something false, since the man whom we claimed to have murdered Smith did not.

A subject-predicate utterance in which the designating term is used attributively may usefully be paraphrased as follows: "The fattest spy (whoever he is) is (must be) a danger." The 'whoever he is' qualifier marks the absence or the irrelevance of any belief held by the speaker as to who in fact instantiates 'the fattest spy'. The 'must be' qualifier indicates the role of a supporting generalization, in this case 'all spies are a danger'. The generalization provides justification for asserting that the fattest spy, about whom the speaker may know nothing, is a danger. He believes that 'the fattest spy' is uniquely instantiated and that whoever instantiates 'the fattest spy' also instantiates 'is a danger'. These two beliefs are the only ones which, should the speaker discover them to be false, would lead him to withdraw his assertion;
they constitute his epistemic ground for referring. These facts suggest that Russell's analysis of denoting describes attributive uses. An utterance in which the designating term is used attributively is equivalent in a very strong sense to the Russellian paraphrase; but we should resist considering the quantificational analogue as an analysis of what happens in the subject-predicate utterance just because their grammatical structures are not mutually reducible and so the semantic functions performed by the parts differ.

I offer the following as the best way to capture the essential difference between the referential and attributive uses:

(1) One uses an expression \( D \) referentially only if there is an object \( x \) such that one intends, by one's use of an expression \( D \), to single it out for one's hearer.

(2) One uses an expression attributively only if one intends by one's utterance of a denoting expression \( D \) to bring it about that there is some object \( x \) singled out for one's hearer.

The position of the existential quantifier outside the scope of 'intends' marks the more intimate relation between intender and object in referential uses. The position of the quantifier inside the scope of 'intends' marks attributive uses and indicates the absence of that more intimate relation between intender and object,
suggesting further that what relation there is depends essentially on which the expression D is.

These two forms have obvious parallels with the exported and non-exported forms of existential belief-ascriptions, e.g. 'there is an object such that S believes it is P' and 'S believes that there is an object which is P'. We use the exported form to indicate that there is some one object to which the believer is related by his belief-state, and we use the non-exported form when we do not want this implication. Given some ascription of belief 'S believes that X is P', we may infer the non-exported existential form 'S believes there is something which is P'. If we may also infer 'there is an object such that S believes that it is P', S's belief-state puts him in a relation to an object, so we say S has a relational belief; while if we may not infer the exported form, S's belief is non-relational. These distinctions apply to intentions also: a relational intention will be one 'directed at' a particular object, and a non-relational one will not be. In Section 5 I attempt to ground the two sorts of intentions on the two sorts of belief as part of a program to explain the epistemic features fundamental to referring.

I am now in a position to deal with a problem left unsolved in Section 2, namely what happens when referring fails. Statements (1) and (2), which I offered as distinguishing the referential and attributive uses, provide the key to formulating the referring
intention in a way which reveals its success-conditions: it will be a relational intention. A speaker has a referring intention, therefore, only if it in fact has an object; this means that the referring intention alone can serve as the antecedent in the 'intend'-to-'refer' conditional.

If there exists an object x, such that in uttering D as part of uttering U, a speaker S intends to bring his hearer H to believe that x is the object to which S attributes some property (or orders something done) in the course of uttering U, intending this to come about in part by H's recognition of S's intention, then S refers to x.

Should S fail to have a referring intention because there is no object for his intention, he will fail to refer; he will still have some sort of intention, of course, but it is not relational. The belief S intends to induce in his hearer, S's intention, and S's belief about X all are relational; the corresponding intention and beliefs for the attributive use are non-relational.

Referring, we have seen, demands a Gricean analysis. Attributive uses do also, but for reasons having nothing to do with their essential features. Donnellan says he thinks the referential/attributive ambiguity is a pragmatic one; I take this to mean that it is not marked syntactically. Referring needed Gricean treatment
we saw, because the referent is identified as the object the speaker intends to pick out; but there is no such object for attributive uses because the speaker's first-level intention for such uses is non-relational. If the attributive use were signalled by a distinctive grammatical form, only first-level speakers' intentions would be relevant; a hearer would not need to recognize which object the speaker means to pick out, and indeed could not, since there is no such object. But just because the distinction is a pragmatic one, the hearer must recognize when a speaker uses an expression attributively, so the use must be characterized by a Gricean intention:

If, in uttering some denoting expression D as part of the whole utterance U, S intends that there exist an object x such that H comes to believe that x has some property P, intending this to come about in part by H's recognition of S's intention, S picks out x attributively.

S determines which object it is that is picked out by his choice of expression, so that he picks out that object, whichever it is, which D denotes, but neither the speaker's intention, the belief which supports it, nor the belief he intends to induce is relational.

5. The Epistemic Foundations

My account of reference irretrievably parts company with the
standard Strawsonian\textsuperscript{9} theory in this section. If I had accepted the standard account, I should say that the knowledge which supports reference to a particular is that which lets a speaker respond explicitly and unambiguously to a question of "Which did you mean?"; he identifies the object for us. We may call this the principle of identifying knowledge. A person has identifying knowledge of a particular, Strawson\textsuperscript{10} says, if and only if he can (i) give a description uniquely identifying of it, or (ii) pick it out from among the other objects in his current field of perception, or (iii) give its name and, while perhaps being unable to give any identifying description except one which incorporates the name, be able to recognize it under a suitably wide range of conditions. I shall shortly present a few examples aimed at showing the principle of identifying knowledge to be false; I don't argue at length against it because Donnellan\textsuperscript{11} and Kripke\textsuperscript{12} have each mounted very convincing arguments against that principle. They claim that proper names

\begin{itemize}
  \item \textsuperscript{9} P. F. Strawson, "On Referring" and "Identifying Reference and Truth-Values", in Strawson, \textit{Logico-Linguistic Papers}, Methuen, 1971, pp. 1 - 29 and 75 - 95, respectively.
  \item \textsuperscript{10} P. F. Strawson, "Identifying Reference and Truth-Values", in Strawson, \textit{Logico-Linguistic Papers}, Methuen, 1971, p. 77.
\end{itemize}
pick out their referents not through associated descriptions and
recognitional abilities, but through a causal connection obtaining
between a speaker's knowledge and use of a name and the person to
whom the speaker learned to apply that name. What is lacking in
their arguments, though, and what I hope to provide, is a satisfy­
ing general account of how it must be with a speaker if he is to be
able to refer. I claim that in order to refer to a particular, one
needs a relational belief about it, and suggest that the preferred
account of relational belief is Kaplan's, which gives a believer's
identifying knowledge of objects no essential role in determining
the objects of his relational beliefs.

(a) Referring

I propose (i) that the object of reference be identified with
the object of the relational intention a speaker has in referring,
and (ii) that the object of a relational intention be identified
with the object of a certain relational belief.

One would expect having an intention to analyze into a complex
of beliefs counterfactual and otherwise, and an action performed
or contemplated. We arguably do have intentions directed at parti­
cular objects, and combining this, which I presume to be a fact
about intentions, with the truism that we cannot have an intention
about or toward X unless we believe that X exists, we should have
an account not of relational intention particularly, but of any

13. David Kaplan, "Quantifying In", in Words and Objections,
sort of object-directed intention. The distinction is easily made, however. If I intend that p, I believe that I shall be able to bring it about that p. For the referring intention, this associated belief is that I shall succeed at singling out X for my hearer. And, in having this belief about the outcome of some action of mine, I also have other beliefs about X, simpler ones of the form 'X is P'. In general, an intention is relational or non-relational depending on whether the corresponding belief is relational or non-relational. This distinction for belief marks certain epistemically important differences; the distinction in belief explains the differences between the two sorts of intention and, through the distinction for intention, the difference between referring and attributing. I suggest that the object of reference is identical with the object of the referring intention, and so with the object of the associated belief.

The belief one has while engaged in contemplating whether to bring it about that, for example, X is Øed, or in trying to do so, is an **occurrent** belief: one comes to have it at some time t, and one ceases to have it at some time fairly shortly after. A speaker's connection with an object he refers to ought to be embedded in his more permanent epistemic condition, though, so the relevant beliefs are those simpler beliefs about X a speaker has which, supported by beliefs about his hearer, lead him to think he will be successful should he attempt to refer. In the ordinary case, when one says
"The Ø is P," one believes that the Ø is P. When we tell someone something, we usually do so because we believe it and desire to convey that information to our hearer. Supplemented with an account of what it is to believe relationally that the Ø is p, this looks like an account of the relevant aspect of the connection between the referring intention, the associated relational belief, and the object of that belief. This can't be right, though, for the simple reason that the speaker may be using 'the Ø' deviantly, and if he is, we will be wrong if we conclude that the object he refers to is the Ø. When a speaker uses 'the Ø' to refer to some object X, he has a relational belief about X, enabling him to have a relational intention directed towards it. From this, however, it does not follow that he is forced to refer to X in any particular way; he may use a term which denotes X but he need not. There will be a relational belief or beliefs about X somewhere in the background, of course, belonging to the speaker's lasting epistemic store, but which they are need not be evident.

This feature of the connection between a relational intention and the belief it presumes causes no practical problems -- my (ii) is as a matter of fact of no help in determining referents, and it couldn't very well be, since discovering the intention is not distinct from discovering the associated belief. We very often have evidence that a speaker has the referring intention, and almost as often we have evidence sufficient to let us determine the referent.
If we should lack that information, we can't tell which is the referent, of course, but we should not expect actually to be able to determine it in each and every case. But there must be a kind of connection between speaker and referent such that should it be broken or should it be non-existent, the speaker when he discovers it to be so will withdraw his claim to have referred, and we, should we discover it to be so, will deny that a speaker refers. It is this connection which makes a speaker's belief relational, namely his being related in some preferred way to the object of his belief.

It was Quine\textsuperscript{14} I believe, who first used the term 'relational belief'; he says that it is the mark of an ascription of a relational belief, that from it we may infer that there exists some object \( x \) such that \( S \) (the subject of the belief) believes it to be \( P \). From ascriptions of non-relational belief, we may infer only that \( S \) believes that there is an \( x \) such that it is \( P \). The exported quantifier marks the relational feature: there is some object to which a believer's state relates him. But not all beliefs which are in some sense 'about' objects are relational, so I shall use the term only when describing the belief-state of one who stands in a certain preferred relation to an object. I think Kaplan's account of the nature of the preferred relation is the best we have. Any account of relational belief must respect certain facts, for instance

\textsuperscript{14} W. V. Quine, "Quantifiers and Propositional Attitudes", in \textit{Reference and Modality}, Ed. L. Linsky, Oxford University Press, 1971, pp. 101 - 111.
that if any beliefs are relational, then certainly one's beliefs about the familiar and domestic objects around one are. Just as clearly, the person who believes that the fattest spy is a danger, because he knows how superlatives work and believes that all spies are a danger, does not have a relational belief about the fattest spy. Kaplan's account handles these facts very well indeed.

He introduces two relations, that of a name's being a name of X (a particular object) for someone (Ralph), and that of a name's being a vivid name for Ralph. Both of these relations are to be distinguished from that of a name's denoting X. (While presenting his theory, I shall follow Kaplan's practice of using 'name' to refer to any designating phrase, and 'denoting' to mean any relation of conventional reference.)

The essential move is to distinguish the descriptive content or import of a name from its genetic character. How Ralph comes to know some name, in the sense of learning to use it to refer to some object, provides a way of relating Ralph's knowledge of and use of the name to an object. The name-of-for relation is a causal one holding between Ralph, the name, and the person or object which stands as the object to which he learned to apply the name. Ralph may have learned the name by having someone introduced to him as N, by being told that the person who he has observed to be the such-and-such is named N, by being present at a dubbing, or by being told about the person named N, in which case the causal element
passes via Ralph's informant. Each of these ways (and some others) puts Ralph in a causal relation to the person to whom Ralph learned to apply the name N. The causal connection, or chain of causal connections, is the genetic character of the name, and is what makes N a name of someone for Ralph.

We need also to separate the descriptive content of a name from some associated descriptive features. A definite description has descriptive content because it contains general terms, and proper names have certain sets of descriptions communally associated with them. This sort of descriptive content is user-independent; the notion we need is not, but is to capture that set of descriptions, beliefs, mental images, etc. which Ralph associates with N. Kaplan calls this the vividness of the name N for Ralph. If Ralph has a relatively detailed, though not necessarily correct, descriptive picture he associates with a name, that name is vivid for him; speaking metaphorically, we may say that Ralph uses the name as a 'label' under which to collect a mass of descriptive material, descriptions, beliefs, and sense-images. The vividness of a name for Ralph depends only on those descriptive features he associates with the name, and not on the genetic features of the name for him nor on the descriptive features communally associated with it. Vividness has no necessary link with the actual world -- rather, it has to do with the purely internal features, as Kaplan puts it, of a name's individuating properties. Ralph may have a 'mental
life' which does not tie very closely at all to reality, but the characters who figure in his mental life may be sharply delineated for all of that. The use of a vivid name makes an existence-claim, but the vividness itself establishes neither that the name denotes nor that it is of someone.

Vividness is explained through the descriptions, names, and mental images a person associates with a name, so if the name is vivid for someone, a complex set of counterfactuals are true of him. For example, if N is a vivid name for Ralph, he would say many things of the form 'N is P' if ... (something). What we might call the propositional part of a name's vividness can be captured with a set of conditionals concerning Ralph's dispositions to use the name. The non-propositional elements can't be so easily or directly put in terms of counterfactuals, though I think it can be done. These other elements are remembered sensory or imagined states of Ralph, and his having them must in part be cashed in terms of what he can recognize, what he can describe, and what he will consider likenesses to the things real or imagined for which he has mental images.

Kaplan's aim was to find a relation holding between a person, a name, and an object (the object of the belief). He wished to characterize those names occurring in belief-ascriptions on which we may existentially generalize, and from which we may infer the form with exported quantifier. We can now define such a relation.
Kaplan calls it 'representation'. He says that a name $N$ represents $X$ for Ralph -- $R(N, X, \text{Ralph})$ -- if and only if (i) $N$ denotes $X$, (ii) $N$ is a name of $X$ for Ralph, and (iii) $N$ is a vivid name for Ralph.

We have this as Kaplan's preferred form for ascribing relational beliefs:

$$ (R): \text{There exists a name } N \text{ such that } R(N, X, \text{Ralph}) \text{ and } \text{Ralph believes } \overline{N \text{ is } P}. $$

Kaplan wanted two things, a rendering of belief which would be transparent to substitution by Leibniz's Law, and a way to characterize those beliefs for which such substitution is legitimate. If there is no object for the belief, then pretty clearly we cannot substitute, while if there is, and if the believer is related to it in the preferred way, we may, so we need to mark off those beliefs which allow inference of the exported existential form. The first conjunct marks the belief as relational since it puts Ralph in a relation to the object $X$. The second conjunct is Kaplan's way of expressing what Ralph believes. Sentences embedded in others in the way that $p$ is embedded in '$S$ believes that $p$' cause frustration because we cannot see how, or if, the sentence which is embedded syntactically is also embedded semantically. Kaplan in effect steps around the problem with his '$B \overline{N \text{ is } P}$' because the words in the sentence inside the square corners refer to themselves, so the whole is in that sense semantically articulated. He views this way of treating belief-ascriptions as part of a program for dealing
with these problematic contexts, namely that we must not suppose that the references of terms in certain sorts of contexts are the same as their references outside such contexts. We shall explain the truth of 'Ralph (B)elieves $\overline{N}$ is $\overline{P}$' partly in terms of Ralph's dispositions to speak and otherwise behave. I shall have nothing to say about the many complex and interesting reasons Kaplan offers for preferring this form, nor shall I offer criticism of it. It seems clear that, whether Kaplan's explanation of relational belief is the best we can do, it won't license false inferences, and it will explain much about beliefs.

I'm interested in just the first conjunct of R, since it tells us how a believer and the object of his relational belief are related. Before I can put this analysis of relational belief to work, though, I need to make certain minor alterations. First of all, I wish to eliminate clause (i) which stipulates that the name N denotes X. Kaplan wanted it for a variety of reasons, amounting to a desire to normalize his theory to the standard form. His discussion of the name-of-for relation and of vividness shows that he views (i) not as a condition on believers but on the formalism -- the connection between a believer and the object of his belief depends not at all on what the denotation of some name is, or on whether it has one. When I use Kaplan's analysis of relational belief to explain reference, I shall be using it without clause (i) therefore. Second, Kaplan thinks that vividness being a matter of degree, there is
some minimum degree of vividness a name may have, below which it doesn't properly distinguish a character in one's 'mental life or inner story'. He realizes that what we take as the minimum degree of vividness is purpose-relative, though, so he allows that we may define senses of 'relational belief' varying just with the degree of vividness to be required of the name. Kaplan appears to think of relational belief in terms of one's having a particular person or thing 'in mind', and so he thinks quite a high degree of vividness is required, but we arguably have beliefs in the preferred sense about persons without having a detailed enough picture to count as having them in mind. I'm inclined to think that degrees of vividness really don't matter to the relationality of a belief, just so long as the relevant name has vividness to some degree and is a name of someone for the believer.

With these few qualifications, I shall put Kaplan's theory of relational belief to work explaining how we are connected with the objects of our references. Things work smoothly if we refer with a name which represents someone for us; it collects a mass of descriptive material around it, and it is the preferred and sometimes the only name we would use to refer to the object it represents for us. We can't knowingly have false beliefs, of course, so if we have a vivid name we might attempt to refer with it. We succeed if the name represents, that is to say, if the name is both vivid and of someone for us; we then have a relational intention directed
at the object represented by the name.

The name-as-label metaphor doesn't really work well except for certain cases, as for instance our uses of 'Homer' and 'Aristotle'. Kaplan uses 'name' for a designating term of any sort, and so it is probably better to give up the metaphor, letting the grouping of descriptive material take care of itself. Our use of names doesn't affect the grouping anyway, and names are not particularly important except for cases like our knowledge of Homer where, if Donnellan and Kripke are right, it is our knowledge of the name itself which, via a causal connection, picks out Homer or Aristotle. In ordinary cases, and particularly in our beliefs about objects of our experience, no single name has dominance over another, and from the body of descriptive material we may extract or create many names, any one of which will be as good as the next for the believer's purposes in referring and for our purposes in ascribing relational beliefs. I think that the dominance of names and the name-as-label metaphor in Kaplan's theory is a matter of expediency; it leads, however, not just to a misplacing of emphasis, but also to a mistake.

I know a person named 'Aristotle', yet I don't confuse him with the philosopher, so if we are to put the causal connection to work, we need to do it in some way other than by saying that one's knowledge and use of the name 'Aristotle' connects one causally with Aristotle. Kaplan at one point seems to suggest that the whole mass of descriptive material grouped around a name is in a sense
the 'real' name. This won't do, though, because it isn't that
group of descriptive material which is the name we use, for example
in referring. We may, however, consider the causal connection to
hold between a certain vividness-group, as we may call it, and the
person or thing to which that vividness-group causally connects.
Vividness-groups are mental states, so they can enter into the
proper causal relations with objects. We shall in this way distin-
guish my uses of the name 'Aristotle' since, depending on the occa-
sion, the use of the name connects with one vividness-group or
another. We may continue to refer to vividness-groups by the names
which serve as 'labels' for them at least in those cases where
the name is dominant, the rest of the material in the vividness-
group serving merely to distinguish that group from others formed
around a similar name. But we ought to separate dominance, inso-
far as it connects with the name-as-label view, from the name-of-
for relation, which actually relates a vividness-group and a person
or object. In most cases the entirety of the descriptive material
in the vividness-group relates causally to an object, and only
when it doesn't shall we want to invoke the notion of dominance in
order to locate the object of the causal connection.

In ordinary cases, the 'dominance' of some one name is illu-
sory, being just a matter of descriptive convenience. Should some-
thing go wrong though, and there be conflicting causal connections
among the descriptions and names in some vividness-group, we shall
look to the internal ordering within the mass of descriptive material: believers are less certain of the fittingness of some parts of that material and more certain of other parts. I may well be unsure what someone's name is, but be utterly certain that a particular visual image I have is a memory of him. We should be unwilling to deny that some vividness-group bears the name-of-for relation to an object when it is faulted in a minor way with perhaps just one deviant element. When more deviations occur though, perhaps when the group divides roughly in half, one half causally connecting with one object and the other half with another, we shall sometimes want to say that the believer does not have a belief about a particular object. Donnellan\textsuperscript{15} discusses a case in which, though a vividness-group relates causally to two persons, we want to use a part of that group in establishing reference. Suppose one knows someone, A, by reputation, knows his written works, etc. and then meets a person, B, whom one believes is A. Later, in talking with a friend, one might say, "A wrote such-and-such a book," and, "A said so-and-so at the party where I met him." In fact it was B who the speaker met at the party, and not A. Donnellan suggests that we shall want to say that in discussing his books, the speaker in fact referred to A, but that in reporting B's remarks, B was referred to rather than A. The speaker, of course, thinks they

are the same person. The basis for so sorting utterances and referents is not clear, but if Donnellan is right, then at least sometimes the reference of an utterance is determined, not by a unique causal connection holding between an object and a vividness-group, but rather by the causal connection holding between some part of the vividness-group and an object. What is unclear is the rationale behind sometimes saying that reference fails because the vividness-group has no unique causal connection with an object, and sometimes saying that it succeeds because some part of the vividness-group does relate uniquely to one object. I conjecture that nothing of theoretical importance is at stake, that in such cases we are applying a principle of generosity, exhibiting our disposition to construe utterances as sensible whenever possible.

We can use Kaplan's theory to explain both what a relational belief is, and how from some one relational belief we can generate many different referring expressions all of which represent the same person or object. Suppose some referring expression N is used with a relational referring intention I. In uttering N, the speaker intended to pick out an object X for his hearer and so he believes that by uttering N he will succeed at doing so. From this complex belief we may extract the simpler belief that X has some property P (unspecified) and in the ordinary case, we may assume that N is one of the names which make up the vividness-group which relates causally to X. In order to cover all cases, though,
I need this extension, that one may use some particular relational belief in support of one's use of a 'dummy' referring term, one which has no place in the vividness-group used to characterize the belief. But if one does have a relational belief about a particular object, one can surely refer to that object with whatever term seems convenient, supposing of course that one's hearer will understand.

The referent has been identified with the object of a certain relational belief, and, given Kaplan's account of relational belief, it follows that a speaker does not need identifying knowledge of an object in order to refer to it. Donnellan's and Kripke's arguments against the principle of identifying knowledge are, I think, strengthened considerably when we see that there really is an alternative to that principle, that we can give it up without giving up hope of getting referents specified in some firm and systematic way.

Donnellan and Kripke rest their arguments on examples concerning the use of proper names to refer to historical persons. They argue, much as does Kaplan though in considerably more detail, that the causal connection is what essentially determines the referent. But if we view the connection between referring term and referent in the way Kaplan's theory suggest, we can extend what Donnellan has called the 'historical explanation' theory of names to descriptions as well. Martin and Deutscher have argued quite convinc-

ingly that in order to have a memory of someone or something, one's memory-state must be causally connected with that thing, via memory-traces connecting with one's original experience of that thing. Consider a memory of some person, which will present itself to the rememberer as a cluster of visual images, remembered states of affairs, etc. That is, it will be a vividness-group which is causally related to its object. I am sure that I have memories of particular persons, clear memories and detailed, but lacking features from which I can construct identifying descriptions. I am sure, for instance, that I have a memory of a particular child in my first grade class at school, since several visual images from that time are still very clear to me, and this child figures centrally in some of them. But I have not succeeded in constructing an identifying description from my memories; they are not detailed in the right way. And certainly I could not recognize that child now, for he would have changed, nor I believe could I recognize him if he were presented to me now looking just as he did then -- how I perceived the world as a child was I think quite different from the way I now perceive it. I submit, though, that I can refer to that child, indeed that I have just been doing so. In one way, there is no problem since, given Martin and Deutscher's claims about memory, I can use as an identifying description 'the child of whom I have these memories'; but one may not know that memories have a causal element, in which case one would not know that
anything was gained by referring via the memory. We may, therefore, construct various non-identifying descriptions and refer with them to persons whom we remember but of whom we have no identifying knowledge. We can construct counter-examples to the principle of identifying knowledge, that is to say, not just with names but with descriptions.

(b) The Attributive Use

The beliefs which support the attributive use of designating expressions are non-relational — they are not in the preferred sense about particular objects. Should I say, "The fattest spy is a danger," I shall almost certainly be using the term attributively. My remark would arise in some context in which, because I believe there are spies, because I believe all spies are a danger, and because I know how superlatives work, I should have occasion to decide that I believe that the fattest spy is a danger. I believe, that is, that 'the fattest spy' is uniquely instantiated, and that whoever instantiates it also instantiates 'is a danger'. Attributive beliefs are a particular kind of non-relational belief; 'the fattest spy' does not represent anyone for me, or as we would commonly say, there is no one in particular who I think is the fattest spy. Russell's analysis of denoting provides, I think, the preferred way to express attributive beliefs. Paraphrasing 'S believes that the Ø is P' we get 'S believes that there is one and only one thing which is Ø and it is P'.
There are two cases to be considered in bringing attributive beliefs in support of attributive uses of designating expressions. One may believe both attributively and relationally that \( N \) is \( P \), or one may have only the attributive belief. My belief about the fattest spy falls into the second class. If one has only an attributive belief that \( N \) is \( P \), one must be using \( N \) attributively in asserting that \( N \) is \( P \). But when one believes both relationally and attributively that \( N \) is \( P \), one may be using \( N \) either referentially or attributively. In order to decide which, we must know what the speaker would count as falsifying his assertion. Suppose, as in a previous example, someone has murdered Smith in a nasty way, and for that reason we believe attributively that Smith's murderer is insane. Now, we might believe relationally of someone that he is Smith's murderer and that he is insane. If, however, I should assert that Smith's murderer is insane, I shall be speaking attributively just if I take the truth of my claim to be independent of the truth of my relational belief about the man I believe murdered Smith, depending instead on two things: (i) that 'Smith's murderer' is instantiated, and (ii) that whoever instantiates 'Smith's murderer' is insane. I will then have asserted something with the force of, "Smith's murderer, whoever, he is, must be insane." By contrast, if on discovering that I believed falsely of \( X \) that he murdered Smith, I withdraw my assertion, I show that I was referring to \( X \).

Only relational beliefs justify the inference of the correspond-
ing exported existential ascription of belief. If I believe relationally that Smith's murderer is insane, it follows that there is someone of whom I believe that he is Smith's murderer and is insane. By contrast, we may infer only the unexported existential form from an attributive belief. From the fact that I believe that the fattest spy is a danger, we may infer only that I believe that there is someone who is such that he is a spy fatter than any other spy, and that he is a danger. The characteristics of attributive beliefs explain non-relational designating intentions. I do not have an intention directed at a particular person or thing when I use an expression attributively, but I can intend to produce a designating expression which denotes, and this is what we do in speaking attributively.
Chapter 4

Referring: The Standard Theory

1. Identifying Knowledge

I shall present briefly what I take to be a widely accepted view of reference. It is a composite theory taken mostly from Strawson\(^1\) and Searle,\(^2\) although I think many other philosophers hold the same or similar views.

A speaker refers in order to identify a particular object for his hearer from the range of objects he believes his hearer already knows or presumes to exist. The speaker wants to talk about some one of those objects, and the function which referring serves and through which it is characterized is that of identifying for a hearer a particular object. A speaker can do this only if he himself has identifying knowledge of the object he means to refer to. He succeeds in what he intends if he picks out a particular object for his hearer by evoking in him, reminding him of, or bringing to mind, the identifying knowledge of the referent he is presumed already to have; he does not communicate that knowledge, since he has presumed his hearer already to have it.

A speaker makes identifying reference to an object \(X\) with an


expression R only if he utters R with the intention of thereby evoking in his hearer identifying knowledge of X. Searle adds an explicitly Gricean element here: the speaker intends to identify X for his hearer by means, in part, of the hearer's recognition of that intention. Strawson's account of reference hasn't a Gricean element; he mentions only the speaker's intention to evoke identifying knowledge. One doesn't know how Strawson would complete his account. He never, so far as I know, does more than discuss certain central and necessary conditions for reference, not attempting to state both necessary and sufficient conditions.

I call Searle's and Strawson's theories 'standard' with respect to the epistemic component, and with respect to the first-level intention a speaker has for R, so differences in how each might go on to complete his theory don't enter. The epistemic component, that is, the role of identifying knowledge, is one of a few aspects of reference about which there has been wide agreement, and for that reason I call theories incorporating it 'standard'.

A person has identifying knowledge of some particular if and only if he

(i) is able to pick a thing out in his current field of perception, or if

(ii) he knows that there is a thing (not in his current field of perception) which is uniquely the such-and-such; that is to say, he has a certain description which he knows applies to one object and no other; such a description is an identifying description, or if

(iii) he knows the name of a thing and is able under a suitably wide range of relatively ordinary or usual conditions, to recognize it when he encounters it, even if he can normally give no identifying description of it other than one which incorporates its own name.

According to the standard theory, then, one has identifying knowledge of an object when and only when one can recognize it, pick it out of one's field of present experience, or give a description which is true uniquely of it. I have argued in Chapter 3 that knowing of some description that it is uniquely true of an object does not count as having a belief about a particular object, and so does not enable one to refer. I illustrated this claim with the following. Someone believes that there is a shortest spy because he believes that there are spies, and because of the character of the predicate 'shortest'. This person reasons that, if there are spies, there is a shortest (or tallest or fattest) spy. He might believe that the shortest spy is a danger because he believes all spies are a danger, and so could say, "The shortest spy is a danger." One who believes that there is a shortest spy does not,
under these circumstances, have a belief about a particular person; he believes of no particular person that he is the shortest spy. One thinks this inference always holds good for reference: if one refers, one has a belief about the particular object to which one refers. If it does, the person in our example did not refer to the shortest spy, but picked him out attributively. The standard theory doesn't distinguish between referring and attributing, and can't do so unless it first distinguishes two sorts of identifying knowledge. I can't, then, really argue that the standard theory's thesis that identifying knowledge is sufficient for reference is false, since the standard theory allows both referring and attributing to count as 'referring'. Referring and attributing ought to be distinguished, though, and when it's done properly, identifying knowledge is found to be sufficient for attributing but not for referring.

Arguing that identifying knowledge is not necessary for reference is difficult, but unlike the above argument, straightforward. Someone who attempts to refer using an expression which is not identifying will fail, according to the standard theory, if the speaker can neither supplement it with a description which is identifying nor recognize the person he means to refer to. But suppose someone has a conscious memory of a person, someone he describes as "an old man from the north of the county"; suppose further that the referrer could not recognize that person if he were presented with
him because the memory is too weak. But he does remember meeting a person whom he knew as the old man from the north of the county. The speaker attempts to refer, and I think he succeeds, since he clearly has a memory of a particular person. In that weak sense he has a particular person in mind. Uniqueness of reference is assured not through uniqueness of description, but through the causal link between the memory and its object.

Martin and Deutscher have argued convincingly that if one has a memory of an object, one's state in having that memory is causally connected, via what they call a memory trace, with one's experience of that object. They argue, further, that what makes a memory one of a particular thing is the causal connection, rather than any descriptive element the memory might contain. One can remember unconsciously, though, and of course one could not refer on the basis of that sort of memory. The standard theory wrong denies that such a memory is epistemologically sufficient for reference; its epistemic part could be re-done, and would become much the same as what I've offered in Chapter 3.

Though in explaining what identifying knowledge consists in, I have paraphrased Strawson, Searle holds similar views. In the next two sections, I look more closely at their individual presentations of theory of reference, since each of them has difficulties

with convention. Searle makes certain mistakes when he attempts to relate the conventional meaning or reference of an expression to how it may be used. How convention functions in picking out a referent is the problem. Strawson’s difficulties also involve convention, but in his case a problem is generated by the conventions of grammar as they relate to what (grammatical) function a speaker intends his words to have.

2. Searle on Reference

Searle’s theory of reference has this as its central element: a speaker S intends that the utterance of R, a referring expression, in the context of a suitable whole utterance, will identify X, a particular object, to a hearer H, and S intends that this come about, in the Gricean way, by H’s recognizing what S intends.

The referring expression R is the means by which X is picked out or identified for the hearer:

"There exists some object X such that either [i] R contains an identifying description of X or [ii] S is able to supplement R with an identifying description of X." 

When (i), R contains an identifying description of X, one sort of connection obtains between R and X. Where (ii), R is not an identifying description (but where the speaker can supplement R with one) another sort of relation obtains. In case (ii), S intends that the utterance of R single out X; the hearer has available,

we presume, and the speaker depends on him having, circumstantial
evidence that X is the object S intends to refer to. Searle is
somewhat careless about this; and as we shall see, he overlooks a
certain connection between what the speaker can intend and what he
can expect his hearer to understand. I let this pass for the
moment, since the problem appears shortly in a stronger form.

Condition (ii) allows that even though R does not contain an
identifying description of X, R may refer to X because the speaker
has a particular object in mind -- an object to which he intends
to refer and of which he can supply an identifying description.
This allows us to refer with a description which is not identifying
so long as we can supplement it with one which is. But it also
allows us to refer with a description which is not merely non-
identifying, but which bears no conventional relation whatever to
the object of reference. Searle clearly wants to allow such cases;
we have dummy referring words like 'thingamajig' which refer to
whatever the speaker intends them to.

Condition (ii) also allows one to flout convention. That is
to say, R may be an identifying description of Y, although one uses
it to refer to X. The easiest to understand of such cases come
about because one knows one is speaking to someone with a false
belief. If you happen to be one of those people who believe Bacon
wrote the Shakespearian plays, and if you suppose I share your
views on Bacon as Shakespeare and know I believe that Lear is the
greatest of the works, you will understand me to refer to it when I say, "Bacon's greatest work is ..." But the conventional referent of that expression is, I think, Novum Organum.

According to Searle's theory, a speaker who refers has identifying knowledge of the object to which he means to refer, and so the referent is fixed in that way. A speaker apparently is free to use a referring expression deviantly because the identity of the referent is always determinate, fixed by the object the speaker means to refer to and of which he has identifying knowledge. Searle explains intending to refer to, or meaning a particular object partially in terms of having identifying knowledge of it; and he explains one's having identifying knowledge in a way which demands of the speaker that he be able to give, on demand, an identifying description of the object.

"... I wish to argue that the two requirements, of uniqueness of intention and ability to identify, are at root identical. ... Clearly the notion of what it is to intend to refer to a particular object forces us back on the notion of identification by description, and we can now generalize this condition as follows: a necessary condition on a speaker's intending to refer to a particular object in the utterance of an expression is the speaker's ability to provide an identifying description of that object."

Condition (ii) allows a speaker to refer to an object other than the conventional referent of the expression he uses, just so long as he can give an identifying description of the object he means to refer to. But in giving the identifying description he

refers to the object — "The one I meant is the φ." — and so we might wonder whether he meant to use 'the φ' deviantly. Certainly a speaker who repeatedly deviated from convention each time he responded to our question, "Which do you meant?" would be acting perversely. The point is that he shouldn't be able to do so; if he can, the referent may never be properly specified — the question, "What is the object of reference?" is just not answered. So long as option (ii) is open, we can pose the sceptical question, "Perhaps he's not referring to the conventional referent of the expression he uses?" Searle has simply not provided us means to answer the question. The way to move against the problem is somehow to limit the speaker's freedom to refer to just any object he might pick using any expression.

The problem comes about because, according to Searle's explicit statement of the rules for referring, one is apparently free to refer with some expression R to an object other than R's conventional referent. But this just cannot in every instance be possible, or else one's hearer can never be sure what one is intending to refer to.! In fact it is not possible, and Searle's rules provide part of the means for showing that it is not. I don't think Searle sees that, though, because the right way to limit a speaker's freedom goes counter to some other views he holds. This has two consequences. First, he doesn't explain to us why it is not possible to refer to just any object one might pick, and second, he can't.
explain how the conventional reference of some expression has any function at all in picking out the referent.

I imagine that Searle thinks he has protected himself against these consequences. As one of the conditions on reference Searle has, "Given that S utters an expression R, ... then in the literal utterance of R, ..."? I take this to mean that a condition on reference is that the utterance of R (as part of a whole utterance) means and is intended to mean what it conventionally does. So, one might wrongly suppose, since the speaker means when he utters R what it conventionally does, he will refer to the object R conventionally picks out as referent.

Meaning and reference just do not connect in this way. When I referred to King Lear as "Bacon's greatest work," it was just because I meant by it what it conventionally means that I succeeded in referring to Lear. Mistaken beliefs generate this sort of case very easily; so the appeal to the literal meaning of the utterance will not help. We are not committed to any particular object of reference just by what our referring expression conventionally means. And anyway, even if this sort of move were to succeed, it would be too strong a limitation on what a speaker can do. One can sometimes refer to an object which is not the conventional referent of the expression one uses even when that expression is an identifying description. On one reading of the conditions (i) and

(ii), reading them with an exclusive 'or', we have have: either R
is an identifying description of some object, or R can be supple-
mented with an identifying description of some object, and not both
R is, and is supplemented, by such a description. This reading
must be wrong because sometimes we can use an identifying descrip-
tion in referring to some object other than its conventional refer-
ent.

I think that Searle does have something like the above in
mind. He includes this clause in his statement of the referring
intention: S intends that his utterance of R will identify X to H
by H's recognition of this intention, and "he intends this recogni-
tion to be achieved by means of H's knowledge of the rules govern-
ing R..." The rules governing R are just those which specify what
it is to intend to refer using R. When one reads the passage I
quoted, one might think that the rule Searle has in mind is one
like "one should use R only to refer to its conventional referent".
But one just cannot get that rule from the ones Searle gives. I
think he believes he includes this rule by the way he phrases his
conditions on referring, and I think he believes that one can't,
except under very special circumstances, deviate from convention.
But this is surely false; we are free to refer to whatever we like
by our words, just so long as our hearer can understand what we say.

This is of course the limitation on our freedom that will fill
the hole in Searle's theory, and disallow us that unacceptable free-
dom we apparently had to refer to whatever we liked by our words. Searle can make the repair quite easily, since it doesn't require a change in his rules, but just an addition.

Donellan\(^8\) has found the solution to this problem; it goes as follows. Generally, we can intend to do something only if we believe we can succeed at it; this is a logical point about intending. Specifically, we can in uttering \(R\) intend thereby to refer to some object \(X\) only if we can reasonably expect our audience to come to know which object we refer to.

This is the necessary condition Searle needs to add to his list. Unless the speaker expects his hearer will understand to which object he refers, he will fail to have the referring intention. In a sense, this condition was already operating since, as I pointed out, unless the speaker expects to succeed at referring, he can't intend to do so. And this means that he really can't intend, perversely, always to refer to some object other than the conventional referent of the expression he uses. Since the source of the limitation is a conceptual truth about intending, someone might say, perhaps Searle presumed it to operate. I think that cannot be defended. A point as important as this about what we can intend is not left unstated; it is not particularly obvious anyway.

And if Searle had presumed it to operate, he would have had two conflicting explanations of how and when one can deviate from convention. I surmised that he believes that we can't deviate from convention in referring except under very special conditions. Although Searle doesn't say what the conditions are, they would probably be like the conditions under which he thinks one can deviate from meaning conventions. As I explained Searle's views on deviating from meaning conventions in Chapter 1, they conflict sharply with Donnellan's view. Searle, I argued, thinks that one can deviate from meaning conventions only when, in some sense, one can replace the proper convention with a different one, as for instance because one's hearer has a false belief about the conventional meaning of some word. One can then speak according to his convention, or his understanding of the convention. According to this view of the matter, one can very seldom deviate from public convention. Whereas, Donnellan's answer allows one to deviate with rather more, though of course not with an unlimited freedom. Searle can't have both principles operating, and I think he would reject Donnellan's.

Donnellan's point about intentions shows how conventions limit our freedom, and when we can deviate from them. If the expression used, R, has a conventional referent, then unless there is something peculiar to the context which gives him reason to believe otherwise, the hearer will take the speaker to be referring
to the conventional referent of \( R \). Since the speaker cannot intend to do something he expects to fail at, he cannot intend to refer unless he expects his hearer will understand what he refers to. The conventional reference of an expression limits a speaker's freedom because it gives the hearer a reason to think that the speaker can deviate from convention only when he can expect to overcome his hearer's standing expectation that he doesn't deviate.

We take each other to be party to the same linguistic conventions; we all speak, say, English. And insofar as we are all party to the same conventions, we individually and collectively have a prima facie compelling reason to assume the other(s) speak in conformity to those conventions. As an application, consider what I write here. I have no reason, and you have no reason to suppose I do have, to be using referring expressions deviantly. And I have a reason, and you know I have, to abide by convention. My aim is clarity and explicitness, and the best way for me to achieve that end is to intend my words to be understood conventionally.

So far as concerns his theory of reference, Searle could accept this revision and use it. But for reasons which came out in Chapter 1, he cannot accept Donnellan's solution. That solution uses the Gricean intention to explain how one's intentions are limited, and accepting Donnellan's solution for referring would put considerable pressure on Searle to accept it for deviating from meaning conventions. And this, as is plain from his Section 6, Chapter 2 in
Speech Acts, Searle will not do. I think that Searle is in the impossible situation of being forced, with reference as with meaning, either to deny that one can ever flout a convention, or to accept the equally absurd alternative that one can mean or refer to just what one wants by one's words.

3. Strawson: A Problem About Grammatical Convention

As is evident in "On Referring" and again much later in "Identifying Reference and Truth-Values", distinguishing referring statements from existence statements is one of Strawson's central and lasting concerns. In that first paper he attempted to show, and to my mind did show, that Russell's analysis of denoting phrases is wrong if it is taken to be an analysis of the correct way to construe ordinary subject-predicate statements. On one version of Russell's analysis, a sentence composed of a denoting phrase and predicates goes into an existentially quantified sentence. In doing so, the referential function disappears or, depending on one's preferences in these matters, is taken over by the bound variables. But whatever else we say about it, Russell's analysis has the consequence that though we thought we were referring to an object with the subject term, we were not, but were in fact asserting its existence and uniqueness. The substance of Strawson's argument against Russell is contained in the role assigned to presuppositions. In referring one presupposes the existence of a such-and-
such; one refers to it but does not assert its existence.

Strawson thinks speakers' intentions crucial: according as the speaker intends to convey to his hearer knowledge of the existence of an object or presupposes his hearer already to know of its existence and intends to make identifying reference to it, he does two clearly distinguished and incompatible things. One cannot have simultaneously the intention to make identifying reference to an object and intend to inform a hearer of its existence — having the one intention precludes having the other. If one intends to make identifying reference one presupposes one's audience to have identifying knowledge of that object. And so, Strawson thinks, one cannot be making an existential assertion using a subject-predicate sentence while referring with the subject term.

Strawson is right that to capture what is essential to the referential intention we must ensure that it is distinct from and incompatible with the intention to assert the existence and uniqueness of an object. His way of achieving this end is faulty in a minor but interesting way.

At a primitive level, and in particular, where no grammatical conventions operate, utterances with structure are differentiated into assertions of existence, and referrings and predicatings by, as Strawson thinks, noting whether the speaker presumes his hearer already to have knowledge of the object in question. If the speaker presumes his hearer already to know of the object, he cannot intend
to inform him of its existence, and so in an utterance which is about that object, we take the speaker to be referring to it. And if he doesn't presume his hearer to know, we suppose he makes an existential assertion. What the speaker presumes his hearer already to know affects decisively what we take his utterance to mean.

Strawson's problem arises because, in a fully developed language, a speaker's intentions can sometimes conflict with grammar. Suppose a speaker utters something in ordinary subject-predicate form; he apparently intends to refer with the subject expression, but he knows his hearer to be ignorant of the existence of the apparent object of reference. Suppose you come up to me and say, "That dent in your new car is going to be expensive to repair," knowing quite well that I didn't know of the dent. This cannot, apparently, be a case of making identifying reference, since you knew I was ignorant of the dent; I had neither identifying knowledge of the dent, nor any presumption of it.

Are we to say that the utterance is faulted in a certain way, relieving us of the need to explain precisely what it means, or what the speaker intended, since it doesn't really mean anything, and the speaker didn't really have coherent intentions? That can hardly be right, since the utterance in my example is perfectly intelligible. It does have an air of perversity, or of a joke, and that is relevant as we shall see.

The utterance in my example is meaningful, about a particular
object, and in the indicative mood. So apparently it must be either an existential statement or a subject-predicate one. If it is the latter, the speaker ought to be referring with the subject term, but if he does, he violates one of the necessary conditions for referring; and the condition he violates is the one which plays an essential role in Strawson's way of distinguishing referrings from existence assertions.

We have a problem about what a speaker could mean in such circumstances. We must be able to say what the speaker meant by his utterance and we can't, according to the Strawsonian theory, say that the speaker meant anything which would require him to be referring with 'the dent in your car'. This precludes us saying that the speaker meant, "The dent in your car is ..." as it would normally be intended, since that would require reference. The existential construal of the utterance, which identifies what the speaker meant with the information which he intended to convey, appears to be the rendering of the speaker's meaning immediately suited to the Strawsonian position. The two apparent possibilities are: "The dent in your car is ..." and "There is a dent in your car which ..." and perhaps Strawson will say that the utterance meant the latter. He may be ready to identify here what one meant by one's utterance with what information one intended to convey, though it seems likely he would draw back from making the identification in the general case. Unless the utterance is judged so
faulted as to be meaningless, I think there is no other alternative for him.

Strawson may find this result congenial; it has some plausibility independently of the good work it does keeping straight referrings and existence assertions. But it just cannot be correct; the speaker is not in this case free to mean by his utterance what we've taken him to mean. I shall argue that the existential construal isn't open, and so that something in the Strawsonian program has gone awry. To do that, I present my objections to the Strawsonian program, and answer the one significant challenge that I'm aware of to my argument against Strawson.

I begin argument with an observation about conventions, since in order for the speaker to intend his utterance to mean, "There is a dent ..." he necessarily intended it in a grammatically deviant fashion. When a speaker utters some sentence, and it is in a certain grammatical form, say subject-predicate, that gives the hearer, if the convention is mutual knowledge between speaker and hearer, a prima facie compelling reason for taking the speaker to mean by his utterance something which is correctly rendered in that grammatical form. For example, if the form of the utterance is subject-predicate the speaker is presumed to refer with the subject-term and predicate with the predicate-term. And if the hearer has no good reason to suppose the speaker to be deviating from convention, the speaker, since normally he will know whether he has given his
hearer such a reason, cannot intend his utterance in a deviant way.
What is questionable is whether, in the circumstance of the dented
car, the hearer had and the speaker knew him to have reason to
suppose the speaker deviated.

There is an objection which could be raised against this line
of argument which, if valid, would provide a competing way of under­
standing what the speaker in our example has done. There is a kind
of deviation from convention which consists in a divergence between
what the speaker intends and what the hearer believes him to intend,
and the speaker might even intend this divergence. The speaker
might use an S-P sentence intending to convey existential informa­
tion with it, while at the same time intending that his hearer
believe him to be using only the standard S-P intention. This
separates what he means from what he intends his hearer to under­
stand. He can't separate them, though, because his intention to
convey existence information cannot in this case be a meaning­
intention, and so he cannot here mean by his utterance, "There
exists a ...". In order to mean that by his utterance, he must
devote from meaning conventions, and since the only grasp we have
of non-conventional meaning is via the Gricean analysis, the non­
Gricean character of our speaker's intention to convey existence
information shows conclusively that he did not mean, "There exists
a ...". The 'divergent intentions' explanation of the situation
won't do. I shall argue that the speaker did indeed intend to
convey existence information, but that precisely because his intention is not a Gricean one, what he means is something else.

We cannot remain on the fence about what the speaker meant. He either meant 'S-P' or 'there exists ...'. If he meant his utterance in the existential way, he deviates from convention, so he must have a suitable Gricean intention, and so he must expect that his hearer will come to know what he intends. He must be giving his hearer reason to think he deviates, since otherwise he can't have the appropriate intention, and the subject-predicate rendering wins by default. If identifying reference were the function conventionally performed by the subject term of an S-P sentence, the hearer would thereby have reason to take the speaker's utterance deviantly, since he knew the speaker knew of his ignorance of the dent. The hearer would then presume the speaker to be making an existential assertion, the speaker could count on the hearer construing him that way, and so he could intend his utterance to mean, "There exists a ...".

In arguing against this conclusion, I want to point out (1) that if the conventional function of the subject term were not identifying reference, but were some less demanding sort of reference, the distinction between referring statements and existence assertions would not be threatened. Strawson implies that it would be, but that's not so. The distinction is firm in the grammar, and the referring function is marked by the presupposition it
carries. And (2), I want to show that only one aspect of identifying reference gives the hearer, or us, reason to suppose the speaker in my example deviates from convention.

The hearer's only reason for supposing the speaker meant his utterance deviantly was his knowledge that the speaker knew him ignorant of the dent. The hearer knows that if one makes identifying reference, one presumes one's hearer already to know of the existence of the object in question. Supposing that identifying reference is the conventional function of subject expressions, our speaker could not make identifying reference to the dent, because he knows his hearer to be ignorant of the dent; and the hearer knows this. Therefore, the hearer reasons, he is doing something else, but what is it? The reasonable guess is that he means "There exists a ...". But whether identifying reference is the sort of referring function which subjects of sentences conventionally perform is one of the things which needs investigating.

I wish to show that there is no general argument, leading from the information the speaker does evidently wish to convey, that there is such a dent, to what the speaker meant by his utterance on this occasion. This bears on both (1) and (2). If we allow ourselves to use, for purposes of argument, a concept of reference less exacting than that of identifying reference, one can, in asserting "The $\phi$ is P," both intend to refer with 'the $\phi$' and intend to inform the hearer of the existence of the $\phi$. To do
this, one depends on the hearer's knowledge of the conventional function (as I suppose it) of the subject term. Its function is to refer and to carry the presupposition of the existence of the thing which is, if not uniquely the $\emptyset$, then referred to uniquely by the speaker's use of 'the $\emptyset$'. Given my supposed conventional function for the subject term, then when you say, "The dent in your car is ..." it is evident which object you refer to. (Strawson agrees that a referring term has the function I give it, though he thinks another convention, the presumption of hearer's knowledge, comes to bear.) So at least sometimes the hearer can come to be informed of the existence of the object of reference, and the speaker can intend to inform him. In this case, the speaker has the Gricean intention with respect to the S-P rendering of his utterance, but he can not with respect to the existential rendering; the utterance can't literally mean two things. The speaker can intend that his hearer infer that there exists a dent in his car, though, and he needn't and oughtn't intend this in the Gricean way.

This shows, I think, that one cannot argue to the necessity of the grammatically deviant rendering from the speaker's evident intention to inform the hearer of the existence of the dent. But how can the speaker intend to refer without flouting the referring conventions which operate in our language? If identifying reference does capture those conventions, I've said, the speaker can't
refer. I point out, though, that the argument has been reduced to a problem about the conventions of a particular language.

If the hearer and the speaker both believe that the referring function in English is that described by the concept of identifying reference, the hearer will have a reason and the speaker may take him to have a reason, to believe that the speaker intends his utterance in a grammatically deviant fashion. The hearer doesn't know of the existence of the dent in his car, and the speaker knows he doesn't, and so the speaker can't make, and the hearer ought not to take him to be making identifying reference to the dent. The hearer still has the problem of discovering how the speaker intends his utterance to be construed, and the likeliest candidate is the existential rendering.

This all seems to work smoothly, but there is a problem with it. The force of the joke is lost, after all, if the speaker is construed simply to be making an existential assertion. And it seems plain that the speaker did mean to make a joke, and to have fun at the hearer's expense. The only way the utterance yields a joke is for the speaker actually to make reference, leaving it to the hearer to infer that there is a dent in his car.

I'm not sure what conclusion one ought to draw, but the one aspect of identifying reference, the presumption of hearer's knowledge, can evidently be violated and yet, for other reasons, we must credit the speaker with referring. We can, of course, qualify
our description of the situation by saying that the speaker's performance is deviant, but in a way which does not grossly fault it. What is unclear is whether we really should say the speaker refers; it seems to me that he did. The consequence of adopting this view is that the presumption of hearer's knowledge is only usually a feature of identifying reference. It is not a feature of reference just when the speaker is sure his hearer will come to know which object is referred to. The sort of joke I've been discussing is one of the few kinds of circumstance in which these conditions are met.

Strawson himself hints that something like this is true. He says, "I cannot claim to be doing more than drawing attention to a characteristic difference of function... For sometimes the operations of supplying such resources [of identifying knowledge] and drawing on them may be conflated in the use of a singular term." But this can only come about by a speaker's referring despite the hearer's ignorance.

I should like to point out that this sort of example forces us to rest the subject-predicate/existential assertion distinction on grounds other than the speaker's presumption of hearer's knowledge: this was my point (1). It is now only a usual or normal condition of reference, and will not serve to distinguish referrings.

and existential assertions. It does mark the distinction when there is no grammatical convention, since in that circumstance one can't make a joke such as the one about the dented fender. One can't play off a grammatical convention and a usual property of reference in the way the joke does, since there is no grammatical convention, and hence nothing to prevent the existential construal; once there is grammar, the case can arise. So it seems to me that we need to depend more on the grammar to distinguish subject-predicate and existential utterances. In particular, an utterance will have the form and meaning grammatical convention gives it unless a speaker can reasonably expect that his hearer would detect him in deviating from the grammar. Only then could an utterance which was apparently subject-predicate in form be really existential in form and import.
Chapter 5
Proper Names

1. Two Theories

When a speaker uses a proper name, how does he succeed in referring to the object named? What is the relation between name and named? Complete answers to these and related questions make up a theory of proper names. What I will call the standard theory of names explains the function of a name by associating with it a set of descriptions. The descriptions carry the burden of picking out a particular object, and the referring function of a name is parasitic on that of descriptions. Names are, in this sense, reduced to descriptions.

Strawson¹ and Searle² have similar theories of names; what I call the standard theory is a composite of their two theories. It is, I think, the most widely held theory of names, and it will be the version I put forward as champion of theories of the type, to face a challenger in the form of a quite radically different theory of names.

1. P. F. Strawson, Individuals, Methuen, 1959; see especially Part One, Section I.

Donnellan, Kaplan, and Kripke have each published papers devoted in whole or part to attacking the standard theory. In the process, a new theory of names is emerging which I call the causal theory. It is by no means fully developed yet, but its general character is clear. I am concerned here to do three things, to show the relation of the new theory to the old, to show that the standard theory is wrong, and to place the new theory of names within a theory of reference.

2. The Standard Theory

Any satisfactory theory of proper names will have it that when we use a proper name, in the ordinary way, we do so with the intent of making known that we are talking about some particular person or thing, and who or what that person or thing is. Our interest in names is part of a more general concern: we need to determine how it is that we succeed in talking about particular persons, how we communicate to a hearer which person we are talking about, and how these two concerns relate.


Both Strawson and Searle exhibit, in their writings on names, the influence of their theories of reference; one's theory of reference should, in part, determine one's theory of names. But inversely, a new theory of names may well require changes in theory of reference. The causal theory does require changes, and earlier in this essay I attempted to develop a theory of reference which would accomodate the new theory of names. Again, Strawson and Searle were taken as representative of those philosophers who defend what I called the standard theory of reference. While difficulties with proper names were not the only source of pressures for revision, the causal theory of names was perhaps the obvious part of theory of reference which could not fit the mold of the old theory.

I concentrate on Strawson's theory of identifying reference because he connects referring to the more general problem of individuating particulars. His theory of reference has a perhaps spurious appearance of being deduced from those other, more fundamental considerations. In *Individuals* for instance, Strawson has larger aims than simply to deal with identifying reference; of those other concerns the only ones I need bother with here are the notions of the identification of particulars, and of identifying knowledge of particulars. One can identify particulars, in the sense that Strawson is concerned with, only when one has identifying knowledge of the particular identified. Strawson explains it as follows.
(i) A man may be able to pick a thing out in his current field of perception, or

(ii) he may know that there is a thing (not in his current field of perception) to which a certain description applies which applies to no other thing: such a description is an identifying description, or

(iii) he may know the name of a thing and be able to recognize it when he encounters it, even if he can normally give no identifying description of it other than one which incorporates its own name. 6

When a speaker makes identifying reference with some expression R, he has identifying knowledge of some object X and intends his utterance of R to identify X for his hearer; that hearer is presumed to have identifying knowledge of X already, so if the speaker's referring expression R is appropriately chosen, it makes known to the hearer that X is the object of reference, rather than informing the hearer of X's existence. Strawson describes this as a speaker's evoking the identifying knowledge the hearer already has.

I want to draw attention to the generality and abstractness of this requirement for making identifying reference: Strawson thinks it is a special case of identifying a particular for someone.

His theory is in a way an *a priori* one: he asks what must be the case for particulars to be individuated, and since he defines identifying reference as the evoking of identifying knowledge, the answer for the special case follows from the general. Why should we assume, though, that all referrings consists in a speaker evoking *identifying knowledge* in his hearer? Existential assertions are paradigms of the conveyance of identifying knowledge, but the speech act of reference may not be so directly connected as Strawson thinks to the general conditions on having and evoking identifying knowledge. (Of course identifying reference is, since it is defined that way.) These more general conditions must be honored by any account of reference, but just how direct the connection between referring and having and evoking identifying knowledge need be, is unclear.

The appearance of the deduction of the conditions on reference from the conditions in general on identifying knowledge of particulars is just an appearance; there is a gap, a break in the reasoning which gives room, as it turns out, for not just a new theory of names but a new theory of reference also. I attempted in Chapter 3 to develop such a theory.

Strawson holds that it is typically the job of the subject term in a subject-predicate sentence to make identifying reference to a particular. Loosely, an expression is used to make identifying reference if it is used to evoke identifying knowledge of a
particular in a hearer. There is some object which the speaker wants to talk about, and the function of the referring expression is to let the hearer know which object that is. So, if the speaker makes identifying reference to an object, (i) he must know which object he means to refer to, i.e., he must have 'identifying knowledge' of the object; and (ii) he must, at least in the standard case, not be intending to give existence-information. In the standard case, the speaker presumes his hearer already to have identifying knowledge of the object of reference, and the function of the referring term is to evoke that knowledge.

It follows, then, that when a name is used to make identifying reference, the speaker using it must have identifying knowledge of the object to which he refers. He must be able to pick the object out which is named, or he must be prepared to substitute an identifying description for the name.

The causal theory of names denies that one must have identifying knowledge of a particular in order to refer to it. One must have some sort of knowledge of the particular, certainly, but it is open to us to argue, and later I do so, that the requirement of identifying knowledge can be evaded in certain cases, and so is too strict. The epistemological foundations of our knowledge of particulars does rest on identifying knowledge. Our practice of referring to objects, though, is much further up in the theoretical structure of our conceptual system; and particularly within
a convention-governed practice of referring, we may expect and do find, that the more fundamental epistemological condition can be evaded.

Strawson thinks a speaker can refer with a name only when he has identifying knowledge of the object he refers to; so the user of the name must be able to give one or more identifying descriptions in explanation of his use of the name. If only one individuating fact is known, a proper name based on that description is just an abbreviation or a stipulative name. Relating a name to a conjunctive set of descriptions is not right either, since we use a proper name somehow to enable us to escape the too great strictness which comes with such a definitional relation.

For instance, we believe that Aristotle was the teacher of Alexander, but if that turned out to be false, our use of the name 'Aristotle' needn't be changed. We would simply report that we were wrong in thinking that Aristotle was the teacher of Alexander. This process cannot extend to all the things we believe true of Aristotle, Strawson and Searle will warn us, since what could we say then, except that we have discovered that Aristotle didn't exist. So the relation between a name, e.g. 'Aristotle' and the facts about Aristotle which are our identifying knowledge of him must be somewhat complicated.

The standard theory relates names and descriptions in this way: with each name there is associated a set of descriptions.
The object named by the name is that object which satisfies a sufficient number of the descriptions associated with the name; in this way we allow for discoveries of historical error without being forced to give up our use of a name. We can still go on talking about Aristotle because, so far as we know, there is an object, Aristotle, which satisfies a sufficient number of the descriptions in the set associated with his name. Of course, if we begin to discover that more and more of the things we thought true of Aristotle were true of no one, or if we discovered that one man was, say, the teacher of Alexander, another the author of *The Metaphysics*, and so on, we would begin to doubt that there was an Aristotle. When a sufficient number of the descriptions associated with a name have proven not to be true of one man, but of no one or of several, we conclude that the name does not refer and that our knowledge of history has been faulty. I purposely leave unspecified what is to count as a sufficient number. Russell thought that all of the associated descriptions had to be satisfied, but that is surely too strong. Strawson and Searle hold only that some, a goodly proportion of, or the like, have to be satisfied.

Strawson thinks the set of descriptions which provide the backing for the name is drawn from the conjoint knowledge of a community of speakers. Any given description in that set is known by some member of the group, but not all need know it. Searle, by contrast, considers the set of descriptions which support a name
to be relative to a speaker. On Searle's view, names have no conventional referents. The possibility of communication is usually preserved, though, since for common names there is considerable overlap between speaker and speaker in the sets of descriptions they associate with a name.

Earlier I said that the standard theory in a certain sense reduces names to descriptions. One must qualify, though, since there is no equivalence between any name and any set of descriptions, and so no real, full-blooded reduction occurs. One sort of discourse is not wholly expressible in terms of another. Searle offers a disjunctive set of descriptions to which, he says, a name is analytically connected. He claims that it is analytic that, if a name names, some disjunction of descriptions is necessarily true of the object named. On his theory of names, it is. But we have not reduced names to descriptions, as Searle recognizes and argues; we see how the two relate by stepping back and talking about theories of names. We require that if a name is to have a referent, a sufficient number of the descriptions in the associated description set must be true of some one object. If the name refers, a sufficient number, but no particular one of the descriptions need be true.

In the standard theory, the referential function of proper names is dependent on the referential function of descriptions in the associated set. If they all fail in any of the ways noted, the
name will fail of reference too. These reductive claims are challenged by the causal theory. In particular, the new theory denies that a proper name must fail of reference if there is no one object which satisfies enough, or even any, descriptions in the associated set; and as a corollary, the new theory denies that a speaker using a proper name must be able to give, on demand, descriptive backing for his use of the name.

3. The Causal Theory

I shall first present the theory of names I call the causal theory, and argue that it is coherent and serviceable. I think it is the correct theory of names, and shall be arguing that it is; but more modestly, I hope to show that there really is a theory to challenge the one I call 'standard'. I shall then discuss a set of examples designed to show that the causal theory is the right one, and that the standard theory is incorrect.

A proper name used designatively will normally be used referentially. A name can be used attributively just when it is explicitly introduced as equivalent to a description; e.g. let us call the shortest spy (whoever he is) 'George'. When we say, "George is a danger," (because we believe all spies are a danger) we are using 'George' attributively because we would be using 'the shortest spy' attributively. One is speaking about that person, whoever he might be, of whom 'the shortest spy' is true.
refers with a name, though, one will intend to pick out some particular object by one's utterance of the name, not depending essentially on the conventional relation of being true-of as a means of fixing the referent. Usually one will use a name to refer to its 'proper' referent, however that is to be explained, since neither with names nor descriptions is it so very often advantageous to speak deviantly.

If one has a relational belief about an object, I've said, then and only then can one refer to it. If a belief is relational, the believer's state in having that belief stands in a certain sort of causal relation to the object of the belief. For instance, I believe that my typewriter sits on a small table in front of me because I experience it to sit there. My experience of the typewriter and the table puts me in a causal relationship with them, and the belief which derives from the experience has a causal relation both to my experiencing the objects to sit so, and to the objects themselves. With every sort of referring expression except names, one's beliefs about the object one refers to have, directly or indirectly, this sort of causal relation to their objects. The direct relation connects one's beliefs with the objects of one's experience, while there are two sorts of indirect relation. In the first, one's beliefs are connected, via one's experience of communication with another person, to his experience of the objects of those beliefs. In the second, one's belief about an object is
connected causally with it via one's experience of some other object whose presence or behavior one knows to be causally connected with the first.

None of these relations between beliefs and their objects work for all uses of names. My beliefs about Aristotle are not essentially connected with any particular experience of mine; I need not remember who told me about Aristotle in order to have a belief about him, whereas I must remember your telling me about the oddly coloured bear you saw on your vacation if I am to have a belief about it. One suspects that the name itself, insofar as it is a part of one's belief, carries the causal connection. Before I try to show how this might work, though, I want to draw an analogy between being a name of and being a picture of. The analogy is from Kaplan and he uses it to much the same purpose as I.

If one knows that a picture is a picture of someone, then one knows something about how the picture came about. If a person figures in the right sort of way as a cause of the picture, then it is a picture of him. Contrasted with the relation of being a picture of is that of being a resemblance of. A picture can be a resemblance of a person without being a picture of him; it can be a resemblance while having originated quite independently of him.

Now, it seems that a description picks out its object in the same way as a picture picks out a person by the relation of resemblance; but a picture of a person picks him out in a different way, in a way, I claim, analogous to the way a name picks out its object.

I prefer to talk about photographs and not painted portraits because the causal relation between the person photographed and the resulting photograph is easier to see than the corresponding relation between sitter and painted portrait. It seems to me that the one who sits for his portrait does figure in a causal way in the origin of the resulting portrait; with a photograph, though, there is a clear causal relation: if the photograph is of someone, then that person was one of the objects which, via light-rays and so forth, caused the photograph to have the patterns of lights and darks it has. This relation will carry through if we take a photograph of a photograph, and etc. At the first step, or at any subsequent step, or link in this chain, a photograph can fail to be a resemblance of the person without failing to be a picture of him. The result of being photographed, as one is aware from personal experience, need not be and often is not a resemblance.

The causal theory of names makes essential use of the causal connection between a speaker's relational belief and the object of the belief. A speaker believes that his use of a name singles out a particular object, and he must have some justification for this belief. His justification is his relational belief that, say, N
is P, where 'N' stands for a name and 'P' for a predicate. Primarily, it is the belief incorporating the name which is causally connected with the person named; and secondarily, a referring use of a name is causally connected with the object of reference via the belief which provides the speaker's epistemic connection with that object. One might compare possessing a belief 'incorporating' a name in subject position with possessing a photograph. The belief can be about a person without the possessor of the belief being able, independently of the causal connection between the belief and its object, to say who the belief is of. Similarly, one can possess a photograph of someone, knowing it to be such, without being able to say who it is of.

We are familiar with the epistemic conditions for referring being expressed in existential form; that is to say, if one intends to refer, one believes that there is one and only one such-and-such, which is the object to which one refers. Such a belief is necessary, but it is not always sufficient, since existential beliefs needn't carry the proper sort of causal relation to their objects. And anyway, if someone believes that, for example, N is P, it follows that he believes that there is something which is uniquely picked out as the object of his belief that N is P. It also follows that there is something which he believes is the object of his belief that N is P, and this exported form of the quantificational version of a belief is tied, so I argued in Chapter 3, to causally backed
beliefs about particulars. I shall continue to use the subject-predicate form to express the epistemic justification one has for referring. The one other sort of belief expressible in subject-predicate form I call the attributive or notional belief; if the believer were to assert what he believes, in having that sort of belief, he would use the subject term attributively.

Names of historical personages will have, according to the theory, a causal chain linking contemporary users of the name, through a sequence of teachers and learners of the name, back through time to the person named. The items in the chain will all be beliefs incorporating the name, and each belief is causally connected with the preceding one via the relation of being learned from one who also held the belief. The person who is at the beginning of the chain must acquire his belief directly, that is to say, through his experience of the person named, or indirectly through experience of that person's causal tracks in the world. If the chain of beliefs is to support reference, none of the beliefs may be faulted with respect to the causal connection; and each of the elements must be a belief. This requirement bars certain cases that a too casual version of a causal theory might allow. For instance, our use of the name 'Santa Claus' relates causally to uses of the name 'St. Nicholas', and there probably was a St. Nicholas. We don't refer to him when we talk about Santa Claus though, because there is no chain of beliefs linking our use of
'Santa Claus' with St. Nicholas, though there is a causal connection. Some of the elements in the chain making up that causal connection will be fictions, and so we don't refer.

A belief-chain connecting our use of a name with the named is the usual case, and nothing epistemically weaker will do. But a belief-chain can be broken, and then later re-established. For instance, if we were to discover that the King Arthur stories were originally told as true accounts of some king, gradually came to be considered fictions, and still later due to historical research, came to be recognized as true, we would begin again to refer to King Arthur, though the belief-chain would have a break in it.

Strictly speaking, there are two belief-chains connected by an underlying causal chain, and by our historical researches: the chain before and the one after the break. So I think we must allow broken belief-chains, so long as the break is supported by (1) an underlying causal chain connecting our use of the name to previous uses, and (2) an appropriate patch in the form of additional information showing that the story we took for fiction is fact. Referring can occur before and after the break but not in the time segment of the causal chain which is the period of the break. We have applied a patch both to the belief-chain and to our theory, but a well-patched chain can be just as good as an undamaged one.

4. Names and Referring

In this section I shall relate what I've so far presented of
the causal theory of names to the explanation of referring I gave in Chapter 3. In Section 5, I return to the causal theory, presenting a set of examples to demonstrate its plausibility and how it might work; from those examples I extract a stronger and a weaker version of the theory.

As with other referring expressions, one can refer with a name without being able to give an identifying description of the person or object to which one refers; conversely, one can have identifying knowledge without being able to refer, since whether the causal connection has an essential role in picking out the referent is just what distinguishes referring from attributing. Simply being able to give an identifying description isn't the appropriate sort of connection, and so one cannot be presumed to be referring. As I shall argue in Section 5, though, the appropriate sort of causal connection is not sufficient to let one refer, since if one's beliefs about the person named are wrong enough, one may not be able to refer. But the connection between the descriptions one has and the object one refers to is not a firm one: the descriptions don't serve to pick out the referent, the causal connection does that.

Suppose one believes that N is P, and one's belief relates to N in the appropriate causal way. Suppose further that N is not P, but is in some sense -- and it may be a rather loose one -- P-like. One may sometimes still succeed in referring to N. The belief which
makes up one's epistemic support for referring can be faulted in the predicate -- the object of reference need not be P -- and it may be faulted in the subject -- the person may not actually be named N -- but it may not ever be faulted in the causal connection which links one's belief that N is P to N. If that connection fails, one does not have a belief about anyone, and so one certainly can't refer.

In the absence of a memory of how it was acquired, and assuming that the causal connection can't be made in any other way, a belief with a description as the subject term will fail to be a relational belief; and if one doesn't know a suitable generality to support an attributive use, one has no reason to suppose it connects with the world at all. This may sound extreme, but in fact it is not; our beliefs about objects form a rather complexly interconnecting set, and no belief which is both about ordinary objects and with which we feel secure fails the requirements I set. I suspect that a belief incorporating 'the supreme being' in subject place might fail these tests, but I don't see that as evidence against them.

Names are different, though; I shall argue that the reason we accept referring with names even when we can't establish a causal connection in the ordinary ways -- through remembered experience -- is because we depend on the conventional reference of the name to fix the referent, and if that conventional relation is a causal one, the causal connection is maintained. This allows names to appear essentially in causally based beliefs. It also explains why
descriptions can't, since the conventional relation they depend on is being true of, forcing one either to speak attributively or to fail to refer.

One sometimes refers with a name in cases in which one hasn't a belief about the referent other than one which essentially incorporates the very name one refers with. Names are unique among referring expressions partly for this reason: they sometimes function essentially in helping to carry the causal connection. Giving the name itself a sometimes essential place in a belief looks rather like giving conventions which govern the name an essential place. If they had, won't one be assimilating at least some uses of names to the attributive use? This doesn't happen, though. First, the only sort of conventional referent-fixing relation which gives rise to the attributive use is being true-of, since no sort of expression except one which picks out its object in that way can figure in a generalization. So, a name can be used attributively only when it is equivalent to some description or set of descriptions. A name's sometimes essential place in a belief serves to individuate that belief and fix its object. When one has acquired a belief about N, but has no conscious memory of a particular experience wherein one acquired the belief, the name figures essentially in carrying the causal connection with the object of the belief.

I want to go back now and consider a certain objection to this theory of names. The causal theory very nearly completely separates
the way a name picks out its referent from the speaker's knowledge of the properties of the object which is named. If it does this, have we not lost our grasp of what it is for a person to be able to individuate an object, to be able to tell it from all others? People who think the standard theory is correct will inevitably bring this charge against a theory like the causal one. I shall cast Searle in the role of one defending the standard theory. He says, "What I am trying to get at is how noises identify objects." His theory of reference is his lengthy answer to this question, and it seems to suggest this explicit short answer. Noises identify objects insofar as they are sequences of words, and sequences of words identify objects insofar as (either directly or indirectly) the sequences of words are true (descriptions) of the objects they identify.

Consider the way Searle handles the 'who?', 'what?', 'which?' questions, answering or forestalling these questions being, he holds, what referring expressions do. He says that the answers to these identity-questions divide into three groups: demonstrative presentations of the object, descriptions in general terms which are true of some object uniquely, and a combination of these two. This account is incomplete. There should be a fourth item on the list, the relation which holds between a word or words and an object

when the words stand in the relation of a conventional referring expression to the object. Strictly, this fourth item includes the relation of being an identifying description of, but what this fourth item includes and what Searle excludes, is the simple case of a word, most plausibly a name, standing as an expression conventionally used to refer to an object simply because a community of language users so use it. Now of course, just because that conventional relation holds, we can give an identifying description of the object of reference: it is the object which is the conventional referent of the referring expression D. We can give the description just because there exists between word and object the conventional relation "X is the object which is named by N." The description depends for its reference on N's conventional reference, and there is circularity only if names depend for their references on descriptions, but they don't.

There is a truth about the necessary place of identifying descriptions, which is that in order for us to state the relation between D and its referent in a non-question begging way, we must be able to refer to D's referent without using D. But we have that means in mentioning D rather than using it: the \( \emptyset \) is the object to which 'the \( \emptyset \)' stands in the conventionally referring relation. This specification of the object is not circular, though we may find it not particularly informative.

But there is a completely different and much more fundamental
way in which identifying descriptions are in fact fundamental. A
version of a Searlian argument might go: I use names to identify
objects, and any way of 'identifying' objects which allows weaker
identifications than as a Ø and uniquely as the Ø will fail for
fundamental epistemological reasons. The objects I can identify
are those I can explicitly identify as, say, Øs; there may be phen­
onema on the epistemological periphery, but they will not be of
central importance. What is right about this is that the practice
of naming rests and must rest on the ability in the strong sense
to identify objects — or else we wouldn't have specific objects to
name. But that the practice rests on this epistemological base
doesn't entail that names and descriptions, in their referring use,
must remain logically tied to that base. There can be and are lin­
guistic practices which break that tie; I have in mind particularly,
connections of utterances of referring expressions with objects
through speakers' memories, and through the causal theory of names.

If this break can occur, the use of names may not be tied in
the way Searle thinks to descriptions. In particular, I argue that
the relation between a name and the object named is not mediated by
and effected through identifying descriptions, and so not by the
relation of being true of. When a speaker refers with a name and
is unable to give an identifying description of the object to which
he intends to refer, he presupposes that there is some object desig­
nated uniquely by his utterance of the name, via either the speaker's
experience or the conventional reference of the name.

Searle's demand for identifying descriptions comes from his
failure, in practice, to distinguish the possibility of giving an identifying description (it must always be possible) and the speaker's ability to give one (he need not always be able).

The new theory of names depends heavily on what we might call the transference of referring credentials. The person or persons who first use the name (the first in the sequence of teachers and learners which terminates with us) must be able to meet stricter epistemic requirements than do those further on in the series. A person who starts a series must have direct causal contact with the person named: he must learn the name directly, either by learning to apply it to a person presented to him, or by learning to apply it to a person whose traces, whose 'tracks' in the world he is directly confronted with. If one decides to call the man who robbed one, 'George', one can initiate a chain of teachers and learners of the name. The use of the name will not rest on the description 'the man who robbed me', but rather on the demonstrative 'the man who did this', pointing to the disturbed furniture, the broken lock, the valuables which are missing, etc. What the demonstrative does that the description cannot do is to tie the reference of 'George' to an event in the world. The causal connection thus established between the future use of 'George' and George via what George did gives the name the proper sort of causal foundation. The effects of a person's agency can go proxy for his physical presence; they are after all the causal marks of his presence, and in that respect like the effects his presence has on our senses.

The individuation of objects is not weakened or even threatened;
the function of a name and its use to refer simply do not carry the 
presumption that thereferrer is able to identify the object of ref-
ERENCE. He has done something which he believes does pick that 
object out, but may not be in a position to check and tell whether 
his referring does indeed do so. He refers taking it on the author-
vity of the person from whom he learned the name that it does pick 
out one person uniquely. His teacher likely learned the name in 
the same indirect way, and so on back through the sequence. What 
matters is that the sequence terminates with someone who can indivi-
duate the person named; but each user of the name needn't be able 
to.

A sophisticated user of the name can give this description:
the person I am referring to by this use of a name is the person to 
whom the name belongs, by whatever means names pick out the objects 
they name. A standard-type theory might disallow this answer as 
question-begging, though it doesn't seem to me they should. It 
should be a legitimate answer, merely making this person's use of 
the name parasitic on the use of the community at large; and they 
must have amongst them descriptive backing of the name. Such an 
answer ought to be allowed by the standard theory. Someone who 
defends the causal theory can also give the above description in 
answer to the question, "To whom do you refer?" He need not be able 
to do so, though -- he refers either way.
5. Application and Argument

I shall offer a rather large set of examples, and they and my discussion of them will, I hope, illustrate how the causal theory as I've so far presented it might work. They also illustrate how one can use a name to refer without being able to give a description which is identifying of some particular object and fixes the reference of the name.

This first example is Donnellan's; it is illustrative of the inessential role of identifying descriptions, and of the special role of names, but it doesn't incorporate any chain of teachers and learners. For this reason it seems to me less controversial and easier to understand.

Suppose a child says to his parents, "Tom is a nice man," and suppose that he can give no further information about Tom. His parents ask questions trying to find out who he is talking about, but the child cannot tell them anything which helps to establish who Tom is. We suppose that the night before the child was awakened by the noise of the party his parents were giving, and came downstairs for a drink of water and a look around. Some guest, Tom, made a fuss over the child, telling him what nice parents he has, what a good boy he is, and so forth. The parents remember that Tom did these things, and conjecture that the child might have

remembered, and so be inclined to say, "Tom is a nice man." We may suppose that Tom is in fact not a nice man; this won't matter. Suppose the child remembers meeting a man named 'Tom', and remembers thinking him a nice man, but nothing else about him -- crucially the child doesn't remember when he met Tom.

The standard theory will deny that the child refers, since according to it, he does not have identifying knowledge of Tom. A defender of that theory might go on to explain this example by invoking the child's memory of Tom as a causal explanation of what he said. That explains what the child did, he might say, but that doesn't make his use of 'Tom' a referring. In the crucial sense, he was not intending to single out a particular, and he could not have been. What is uncontested is that the child's experience the night before of the man Tom is the cause of his saying what he did the next morning; and it is also uncontested that the child's use of the name 'Tom' is causally connected with Tom. What is contested is whether the child has referred.

Now according to the causal theory of names, the child has referred. His use of the name 'Tom' is causally connected with the man whom he remembers; he remembers that the man seemed to him nice and that his name was Tom. The memory is what is crucial: without it we cannot understand at all why the child should have said what he did. He says, "Tom is a nice man," because he remembers meeting Tom and thinking him nice, and the memory allows the child to refer.
One can construct a related example that doesn't play on the lack of intellectual development of a child, but which makes use of ordinary failure of memory -- the child could have said, had he understood what was wanted, that he remembered Tom from sometime in his immediate past. Suppose that someone is telling a story. He says "... and then I met John Smith, who was a strange old man." If the story-teller can give no description of John Smith which is identifying, and if he could not now recognize him, he cannot, according to the standard theory, refer. Even though he knows some things -- about John Smith, one wants to say -- they alone on no theory of reference will allow him to refer, since they do not count as identifying descriptions. The story-teller, we may suppose, remembers some things about John Smith, and can place his meeting with Smith in a certain period of his life, say of when he lived in a certain place, but no more precisely than that. The description 'the man named John Smith whom I met when I lived in such-and-such a place' is available, and we may suppose the speaker to remember this much about his acquaintance with Smith. This description might seem question-begging if used as an identifying description; it makes essential use of the name, and without using it, the description is not identifying. We don't need to settle that point, since we can simply suppose that at that time of his life, the speaker had actually met two men named John Smith, one he knew fairly well, worked with, etc., and the other he met just once, and immediately forgot.
The John Smith that he knew well is the one he remembers, and is the one about whom he is telling the story. The description incorporating the name is no longer identifying, but the speaker has a memory of the one man, and not the other, and not both.

According to the standard theory, the speaker is not intending to refer to a particular person, since he cannot give an identifying description, nor could he, we suppose, recognize someone now as John Smith. His memory of Smith is too vague for that. The standard theory seems in this case very counter-intuitive. I think the story teller has referred to one man, John Smith, just on the basis of his memory of Smith. The content of that memory does not establish the connection, though the memory cannot be contentless, since if one remembers nothing, one has no memory; the connection with John Smith is established by the causal connection between the memory the speaker has now and the man who was the object of the experiences which stand as a cause of the memory.

There is an appearance of infelicity about these example, I think, and it comes about because of our suspicion that the speaker shouldn't have expected his hearers to know which object was referred to. It is important that in each example a name is the referring expression; we seem to accept referrings with names even when, in some sense, the hearer isn't able to tell who is being picked out. If in each of these examples the referring expression had been 'the person of whom I have this memory' we would know that the speaker
had unreasonable expectations about what his hearers would understand — a memory is too much a private matter. Because the referring expressions were names, I think we accept these cases, though it's hard to see where the difference lies, since the epistemic ground is in each case the same. I shall attempt to explain this.

Kripke thinks that if one knows someone's name, one knows who he is.¹⁰ I really haven't much of an idea what Kripke might mean by his claim; it certainly needs explaining. I do think it connects with this point, though: we seem to accept referring using names when the name plays an essential role in picking out the referent, so that in a sense all one needs to know (as opposed to believe) about a person in order to refer to him is his name. I claim that a name has a sort of special position — that when one knows someone's name because one has a relational belief about him which essentially incorporates the name, one can refer even though the only way in which one makes it known to one's hearer to whom one refers is simply to transfer to that hearer one's own knowledge of the name — or rather, the belief which essentially contains the name. This is definitely a weaker claim than Kripke's, though it may be all he had in mind.

Names are something special, I think; there are conventions

about their use which are peculiar to them, not shared by other referring expressions. (I don't want to raise problems about identifying a referring term as a name; I merely suppose that we can usually recognize a name when we meet one.) A name is conventionally used to refer to the bearer of the name. If as a hearer one knows that some term, as it is used in an utterance, is a name, then one presumes that, however names are connected with their objects, this term is also connected with its object in that way. When a conventional connection is present, it gives one a good reason though not an absolutely conclusive one, to think that the speaker refers to the conventional referent. Names and descriptions have conventional referents; and usually they are used to refer to those items. So if we can find out what the relation of conventional reference is for names, we may have a means of showing why my examples seem all right, though strange.

If names connect with their objects in the way the causal theory says they do, and if the hearer supposes them to, then he will know that the object referred to is the object to which the speaker's use of the name is related via a causal connection to its object. If the hearer doesn't believe that names have causal connections, but believes merely that they are somehow conventionally connected with their objects, he will suppose that this use of the name has that conventional relation, whatever it is, to its object also. We can assume a little more about the hearer, namely that
he knows how we use names. He may not be able to tell us what it is, but he will accept or reject referrings partly on that basis.

This brings me back to the point I mentioned earlier, that in practice we accept knowing someone's name, in the context of having a belief about him, as sufficient to let us, as speakers, refer, and as hearers to judge that such a referring picks out a particular object for us. Something very like this works for descriptions, too, with a symmetry on the hearer's part between the two cases, but not on the speaker's part. I can't refer with a description unless I have a belief with a causal base about the object I believe to be, say, the $\emptyset$, but I can pick that object out for a hearer by communicating nothing more to him than that there is a thing which is the $\emptyset$, and it is, say, P. He presumes my referring credentials are in order, and accepts that a particular object, the $\emptyset$, has been picked out for him by me. He now knows there is an object which is the $\emptyset$, and he has a causal connection with it, since I have transferred to him the causal base of my own epistemic connection with it.

There is, though, another way in which this parallel between names and descriptions breaks down: we use names to refer even if we don't remember the occasions on which we learned their use from another person, and even though their reference depends on that other person. Whereas, with descriptions, if the causal base of one's referring use of some description rests completely on a communication from another person, one may not forget how one learned
that description and still use it to refer. I believe that it is unique to names that referring credentials for their use can be passed on in this way. If the causal theory is true of names, it would explain this difference between names and descriptions; the conventional relation between an identifying description and its object is truth, so if the speaker through forgetfulness no longer has a causal connection with the object of reference, he must fall back on the conventional connection which is not causal, and so he can't refer. The conventional connection between a name and its referent is, according to the causal theory, a causal one, so one preserves the causal element even while depending on convention. And so too, one avoids turning one's referring into an attributing.

What I need now is some evidence that our use of names is such as to require description using a causal connection, so I shall discuss examples which tend to show that our use of names for historical persons -- and so persons of whom we have no memories -- fits the causal theory better than the standard one.

All we know about Homer, or at least all I know about him, is that he is the author of *The Iliad* and *The Odyssey*, that he was a Greek, and that he was blind. Now let us suppose that the man who composed those epic poems was not called Homer, but had some other name, and let us suppose that the name 'Homer' came to be attached to the works in something like the following way. The author was obscure, even in his own time, and wanted to stay that way. But
he had a friend named Homer who took the works and brought them to the attention of some local ruler. By bringing them to the attention of a wider public, and by being their promoter, as it were, the works became known as Homer's *Iliad* and *Odyssey*. Homer was honest, we may suppose, and didn't claim them as his own, so no lying was involved. The possessive 'Homer's' was just a convenient way the people of the time had of associating the works and their promoter.

But who do we refer to when we talk about Homer? According to the standard theory, we refer to the writer of the works, since the reference of the name 'Homer' is fixed for us by the uniquely identifying description 'the author of *The Iliad* and *The Odyssey*'. Yet the story I've just made up could be true. Given this story, though, wouldn't we be ready to say that we had been mistaken about what Homer did? I have a parallel case in mind. America was named after Amerigo Vespucci, and not after Columbus; a promoter is often more intimately associated in the public's mind with a thing than the discoverer, inventor, or maker. If it had happened, and for all I know it may have, among at least some people of the time, that they mistakenly thought that Vespucci had discovered America, they would simply have had a false belief about him. The reference of the name 'Vespucci' would not have been fixed by its connection with 'the discoverer of America'. And if we go on to suppose that our knowledge of Colomus, Vespucci, and their times became as
thin as our knowledge of Homer and his times, I see no reason why we should alter our conclusion, that they and hence we, had a false belief about Vespucci.

Donnellan has an analogous example using Shakespeare and Bacon as characters. One supposes that the Shakespearian plays were not written by Shakespeare, but by Bacon; Shakespeare, so the story goes, was just an actor, and not as he in fact was, the author of the Shakespearian plays. The references of 'Shakespeare' and of 'the author of the Shakespearian plays' would diverge, and it is not clear what the standard theory would have to say about the situation. One's intuitions, or at least mine, indicate that the reference of 'Shakespeare' is unaffected, that it goes with the man, whereas the reference of 'the author of the Shakespearian plays' reverts to Bacon. The standard theory might have trouble with this example just because it is easy to think that the most important thing, for us, about Shakespeare is that he wrote the Shakespearian plays, and so the reference of 'Shakespeare' ought to go with the plays. We do have some historical knowledge of Shakespeare which is not essentially connected with his being the author of the plays, and it is that knowledge which dominates in fixing the reference of 'Shakespeare'; but it is not obvious why, attempting to use the standard theory, it should.

The following example is a variation on one of Donnellan's. Thales reputedly said that everything is water. What knowledge we
have of Thales comes from mention of him in Aristotle, Plato and Herodotus, and since descriptions like 'the ancient Greek philosopher who said that everything is water' may not be identifying, we depend quite a lot on descriptions like 'the ancient philosopher to whom Plato and Aristotle attribute the view that everything is water'. I think we give such descriptions dominance over the simpler one for another reason also, that we may not be so concerned with the man who actually did hold that view as with the man to whom Plato and Aristotle attribute the view. This shows something about the causal theory: causal considerations, entering via the chain of teachers and learners of the name 'Thales', are here determining (in the standard theory) which descriptions are important.

Suppose that Thales was not a philosopher but a well-digger, and that Plato found him a congenial drinking companion. Perhaps no one actually held the view that everything is water, but that in his cups Thales once said that he wished that everything was water so that he wouldn't have to dig those damned wells. The use of the name in Plato's writings originates, as a joke, with this man; but he is not a philosopher who held that everything is water. It seems to me clear, though, that the referent of the name 'Thales' is that man Plato knew, the well-digger.

I think these examples show that a name's referent is determined by how the causal chain of learners and teachers of the name began. Someone first learned that some name applied to a particular
person, and either through that first person or through several, the name has been passed on from teacher to learner, and so to us. We traditionally learn about Thales in first year philosophy, where we learn that some ancient Greek named Thales had once said that everything is water. What is important is the sequence of learnings of the name, and not the properties attributed to the named person; in the example, Thales wasn't a philosopher and all he said was that he wished that everything were made of water.

The debate whether King Arthur actually existed is similar. We don't believe that Arthur existed, so we don't attempt to refer to him except as a fictional character, and so we could not discover that we had been referring to him. There are clearly interesting parallels, though, between the problem of determining who a story is about, and who we refer to; the latter seems to be a kind of special case of the former. No one expects that there was a king who did in detail the things Mallory tells of Arthur. That is not the point, and that is not what is in question when one wonders whether Arthur existed. Rather, one wants to know whether there was some historical personage, a ruler of some sort, who in some way gave rise to the Authurian legends.

One can't directly apply the requirement that the person, if there is one, who was Arthur is someone of whom a sufficient number of things commonly attributed to him are true. A standard theorist could begin, though, by finding whether a sufficient number of
Mallory's stories were nearly enough true. One needn't expect any of them to be true; all could have been changed, exaggerated, embellished, in the telling, and if we could still somehow trace the origins of the story back to some ruler in ancient England, we would agree that Arthur did actually exist. This example at least favors the dominance of the causal element in determining who, if anyone, a story is about. In order for a standard theorist to even start, he has to accept non-identifying descriptions. And then even if he finds, through historical investigation, descriptions identifying of Arthur, how does he get from them back to the non-identifying ones making up our present knowledge? It might be possible, but not without, I suspect, a causal element slipping in.

The Robin Hood story is another nice example. We needn't expect that anything of the Robin Hood legends be true in order for the story to be about a particular person. We don't believe that if he did exist, he robbed from the rich and gave to the poor; more likely he robbed and kept. I think that generally it is held that the Robin Hood legend is just that, and that there was no person who in some sense the legend is about -- but there could well have been. I believe that the origin of the legend is to be traced to the time and circumstance of the Norman conquest of England, and that the legend is supposed to have been wish-fulfillment on the part of the conquered Anglo-Saxons; they wanted to get their own back. Now pretty clearly the legend could have arisen about some
leader of a small band who used resistance to the Normans as an occasion for banditry.

I began the argument for the causal theory of names by claiming that some causal connection holds between the use of a referring expression and the object it refers to. This, I claim, distinguishes properly referential uses from ones which are not really referential but something else. Then I presented some examples aimed at showing that sometimes a speaker's memory (a type of causal connection) of some object is sufficient to let him refer. The last set of examples was designed to show that a sequence of causally connected beliefs held by teachers and learners of them, and involving essentially some name, provides a link between name and named when the sequence terminates with someone who learned the name and its use directly. Any one of the persons in that chain can refer without being able to give any identifying description of the object named.

6. A Strong and a Weak Version

What I haven't discussed, and what I don't know how to treat conclusively is the difficulty about just how misinformed we can be about someone and still refer to him by name while depending on a causal connection to achieve reference. Consider again the problem of whether King Arthur existed. Most of the things attributed to him by Mallory might be false, and we would still be right in
saying that Arthur existed just so long as two things hold. Our use of 'King Arthur', that is to say, the Arthurian legends, must have their origin in some actual person, and that person must in some sense be a suitable object. This second requirement is what gives trouble, since we don't know what is to count as suitable. Of course, this requirement is a reversion to a theory of names which gives a role to descriptions in fixing the referent. It seems to me improbable that we would say Arthur existed if the person who did stand at the origin of the Arthur stories was just an ordinary person, having no role as king or chieftain. We might, for instance, be prepared to say that Arthur existed even if it turned out that the legend was imported, and so was not about some ruler in England, but I doubt that we would unless there was both some person who stands as the original of Arthur and who filled the right sort of place in history. Against this, the example I presented of Thales seems to indicate that almost nothing of the things we commonly think true of Thales need to be true so long as the historical connection between our use of 'Thales' and some one person holds.

Perhaps I've made this contrast stronger than it ought to be. Perhaps we can't judge properly whether we would say of someone that he was Arthur even though he wasn't a king or chieftain, unless we know something more about the connection between him and our use of the name. It is logically possible, one supposes, that the
person who stands as the original of Arthur was some unknown peasant or yeoman farmer, with nothing kingly about him whatever. This won't do, though, as a counter-example. What one expects to happen is this. If the story is too implausible, if the successive changes the story undergoes are too improbable, we will say that one version of the story inspired the other, but that they are not both about the same person. One story merely inspired the other, perhaps, and in the transition aboutness has been lost. We need to know more about possible criteria for judging aboutness.

If someone succeeds in telling a really ingenious story which successfully connects our use of 'Arthur' with some peasant in pre-Norman England, we have several options. First, we need to notice that we are not forced to do anything, since we have not been referring in our use of 'Arthur' in the stories, so we can conclude that some person did stand as the original in the Arthur stories without having to affirm that Arthur existed, and so that the stories are about him. This option seems to indicate that we do require that any candidate for Arthur play the right sort of role in history.

If we believed that we had been referring, there would be considerable pressure on us to say that the person who stands as the original of the story is the one we refer to. But if we believe that such a person exists, our epistemic justification for referring will put further restraints on the sort of story that can be told. The story will then have to preserve the causal connection between
the person and our use of the name, and will at the same time have to explain how we now come to believe that that person did exist. This restraint blocks some of the wilder stories, I think, by making it very likely that there will be some sort of relevant similarity between the properties of the original and the properties we presently attribute to him. In my example, Thales said that he wished that the world was made of water; we can understand how this might have become what we now attribute to him, and that similarity is enough to make us feel comfortable, or so I think, in agreeing that this man, as my story has him, could have been the original of our use of 'Thales' and be the person we refer to when we talk about Thales.

I don’t know, though, that these restraints will make really wild stories impossible. They will be harder to construct, certainly, but someone may succeed in inventing one. Given a long enough chain, we could probably connect a name with any object we might pick, e.g. 'Homer' with the Empress Theodora's lap dog. If that happened, it seems to me that we would refuse to say that we had been referring, and would refuse to say that N existed. We could impose some maximum on the length of acceptable stories connecting our use of the name and its object, but that may not be enough. And we anyway have grounds for thinking it insufficient for other reasons. I think we need to require that there be some property of the original of the story which has some suitable
relation to the properties we think he had. We don't need to require, and we shouldn't, that the original of the story have at least some of the properties we do presently attribute to him -- that is too strong. What matters is that in some sense the person who is the original of the story fills something like the role in history we attribute to him. This requirement is vague, I think. I don't see that there is any kind of sharp line to be drawn. The wilder the stories are, the nearer they are to failures at showing that reference takes place, and at some point, we refuse to believe that we have been referring. But there need be no clear line, and no sharply enunciable criterion.

I haven't shown that the causal theory of names is the theory which describes the way we use names. I have searched for, but so far have been unable to find, a case where we have discovered of some historical personage that all of what we believed true of him was false, yet we continued to refer to him, though changing our minds as to what he was like. Such a case would show that at least sometimes the reference of a name is secure even though its descriptive backing changes completely. The examples and the outline of theory that I've offered do show, I think, that the causal theory is perfectly coherent; we could use names that way without any bad consequences.

And I think I have shown that the standard theory of names is wrong. First, I've argued that, for all referring, there must be a
causal connection between the referent and the speaker's epistemic justification for referring. For names, this requirement most plausibly is that the chain of teachers and learners of the name terminate in someone who had a suitably direct causal connection with the person referred to. This shows that knowledge of identifying descriptions is not sufficient to let one refer, and so is not sufficient to let one refer with a name. And I think that my examples are persuasive of the following: that one need not know any identifying descriptions in order to refer, using either names or other referring expressions. To this extent, and for this much of the causal theory of names, I believe that I have shown it to be the correct theory.

The strong version of the causal theory would deny that descriptions are in any way essential to referring with a name. A weaker version would give them this place: there must be a causal connection between one's use of the name and the person named, but some of the things we believe true of the person named must be either true of him, or sufficiently similar to things which are true of him if the name is to refer. A name and its associated description set 'fits' an historical person when the place in history attributed by the descriptions is filled, more or less satisfactorily, by the object picked out causally. For instance, this requires that Arthur have been a leader of some sort, that Thales be somehow associated with the claim that everything is water, and
that Robin Hood be a leader of a band of (merry?) men.

This set of descriptions does not pick out the object of reference; the causal connection does that. The set serves to delineate, within broader or narrower limits, but never uniquely, the kind of thing the object of reference is. If the causal connection comes in conflict with the description-set, we can emphasize one side or the other, depending on the individual case. But the description-set can sometimes override the force of the causal connection, and in such a case, even though uniqueness of causal connection holds, the description-set, being uniformly inappropriate to that person, cancels reference. It seems to me that this weaker version of the causal theory is the one which describes our practice.
Bibliographical References


*Individuals*, Methuen, 1959.