

PHILOSOPHY OF
MIND

Classical and
Contemporary Readings

David J. Chalmers

New York Oxford
OXFORD UNIVERSITY PRESS
2002

Oxford University Press

Oxford New York

Auckland Bangkok Buenos Aires Cape Town Chennai

Dar es Salaam Delhi Hong Kong Istanbul Karachi Kolkata

Kuala Lumpur Madrid Melbourne Mexico City Mumbai Nairobi

São Paulo Shanghai Singapore Taipei Tokyo Toronto

and an associated company in Berlin

Copyright © 2002 by Oxford University Press, Inc.

Published by Oxford University Press, Inc.

198 Madison Avenue, New York, New York, 10016

<http://www.oup-usa.org>

Oxford is a registered trademark of Oxford University Press

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior permission of Oxford University Press.

Library of Congress Cataloging-in-Publication Data

Chalmers, David John, 1966–

Philosophy of mind : classical and contemporary readings / David J. Chalmers.

p. cm.

Includes bibliographical references.

ISBN-13 978-0-19-514581-6

ISBN 0-19-514580-1 (hardback : alk. Paper)—ISBN 0-19-514581-X (pbk : alk.paper)

I. Philosophy of mind. I. Title.

BD418.3 .C435 2002

128'.2—dc21

2002072403

Printing number: 9 8 7 6 5

Printed in the United States of America
on acid-free papers

C. Internalism and Externalism

The Meaning of "Meaning"

Hilary Putnam

Meaning and Extension

... Since the Middle Ages at least, writers on the theory of meaning have purported to discover an ambiguity in the ordinary concept of meaning, and have introduced a pair of terms—*extension* and *intension*, or *Sinn* and *Bedeutung*, or whatever—to disambiguate the notion. The *extension* of a term, in customary logical parlance, is simply the set of things the term is true of. Thus, "rabbit," in its most common English sense, is true of all and only rabbits, so the extension of "rabbit" is precisely the set of rabbits. Even this notion—and it is the *least* problematical notion in this cloudy subject—has its problems, however. Apart from problems it inherits from its parent notion of *truth*, the foregoing example of "rabbit" in its most common English sense illustrates one such problem: strictly speaking, it is not a term, but an ordered pair consisting of a term and a "sense" (or an occasion of use, or something else that distinguishes a term in one sense from the same term used in a different sense) that has an extension. Another problem is this: a "set," in the mathematical sense, is a "yes-no" object; any given object either definitely belongs to *S* or definitely does not belong to *S*, if *S* is a set. But words in a natural language are not generally "yes-no": there are things of which the description "tree" is clearly true and things of which the description "tree" is clearly false, to be sure, but there are a host of borderline cases. Worse, the line between the clear cases and the borderline cases is itself fuzzy. Thus the idealization involved in the notion of *extension*—the idealization involved in supposing that there is such a thing as the set of things of which the term "tree" is true—is actually very severe.

Recently some mathematicians have investi-

gated the notion of a *fuzzy set*—that is, of an object to which other things belong or do not belong with a given probability or to a given degree, rather than belong "yes-no." If one really wanted to formalize the notion of extension as applied to terms in a natural language, it would be necessary to employ "fuzzy sets" or something similar rather than sets in the classical sense.

The problem of a word's having more than one sense is standardly handled by treating each of the senses as a different word (or rather, by treating the word as if it carried invisible subscripts, thus: "rabbit₁"—animal of a certain kind; "rabbit₂"—coward; and as if "rabbit₁" and "rabbit₂" or whatever were different words entirely). This again involves two very severe idealizations (at least two, that is): supposing that words have discretely many senses, and supposing that the entire repertoire of senses is fixed once and for all. Paul Ziff has recently investigated the extent to which both of these suppositions distort the actual situation in natural language;¹ nevertheless, we will continue to make these idealizations here.

Now consider the compound terms "creature with a heart" and "creature with a kidney." Assuming that every creature with a heart possesses a kidney and vice versa, the extension of these two terms is exactly the same. But they obviously differ in meaning. Supposing that there is a sense of "meaning" in which meaning = extension, there must be another sense of "meaning" in which the meaning of a term is not its extension but something else, say the "concept" associated with the term. Let us call this "something else" the *intension* of the term. The concept of a creature with a heart is clearly a different concept from the concept of a crea-

ture with a kidney. Thus the two terms have different intension. When we say they have different “meaning,” meaning = intension.

Intension and Extension

Something like the preceding paragraph appears in every standard exposition of the notions “intension” and “extension.” But it is not at all satisfactory. Why it is not satisfactory is, in a sense, the burden of this entire essay. But some points can be made at the very outset: first of all, what evidence is there that “extension” is a sense of the word “meaning”? The canonical explanation of the notions “intension” and “extension” is very much like “in one sense ‘meaning’ means *extension* and in the other sense ‘meaning’ means *meaning*.” The fact is that while the notion of “extension” is made quite precise, relative to the fundamental logical notion of *truth* (and under the severe idealizations remarked above), the notion of intension is made no more precise than the vague (and, as we shall see, misleading) notion “concept.” It is as if someone explained the notion “probability” by saying: “in one sense ‘probability’ means frequency, and in the other sense it means *propensity*.” “Probability” *never* means “frequency,” and “propensity” is at least as unclear as “probability.”

Unclear as it is, the traditional doctrine that the notion “meaning” possesses the extension/intension ambiguity has certain typical consequences. Most traditional philosophers thought of concepts as something *mental*. Thus the doctrine that the meaning of a term (the meaning “in the sense of intension,” that is) is a concept carried the implication that meanings are mental entities. Frege and more recently Carnap and his followers, however, rebelled against this “psychologism,” as they termed it. Feeling that meanings are *public* property—that the *same* meaning can be “grasped” by more than one person and by persons at different times—they identified concepts (and hence “intensions” or meanings) with abstract entities rather than mental entities. However, “grasping” these abstract entities was still an individual psychological act. None of these philosophers doubted that understanding a word (knowing its intension) was just a matter of being in a certain psychological state (somewhat in the way in which knowing how to factor numbers in one’s head is just a matter of being in a certain very complex psychological state).

Second, the timeworn example of the two terms “creature with a kidney” and “creature with a heart” does show that two terms can have the same extension and yet differ in intension. But it was taken to be obvious that the reverse is impossible: two terms cannot differ in extension and have the same intension. Interestingly, no argument for this impossibility was ever offered. Probably it reflects the tradition of the ancient and medieval philosophers who assumed that the concept corresponding to a term was just a conjunction of predicates, and hence that the concept corresponding to a term must *always* provide a necessary and sufficient condition for falling into the extension of the term.² For philosophers like Carnap, who accepted the verifiability theory of meaning, the concept corresponding to a term provided (in the ideal case, where the term had “complete meaning”) a *criterion* for belonging to the extension (not just in the sense of “necessary and sufficient condition,” but in the strong sense of *way of recognizing* if a given thing falls into the extension or not). Thus these positivistic philosophers were perfectly happy to retain the traditional view on this point. So, theory of meaning came to rest on two unchallenged assumptions:

- I. That knowing the meaning of a term is just a matter of being in a certain psychological state (in the sense of “psychological state,” in which states of memory and psychological dispositions are “psychological states”); no one thought that knowing the meaning of a word was a continuous state of consciousness, of course).
- II. That the meaning of a term (in the sense of “intension”) determines its extension (in the sense that sameness of intension entails sameness of extension).

I shall argue that these two assumptions are not jointly satisfied by *any* notion, let alone any notion of meaning. The traditional concept of meaning is a concept which rests on a false theory.

“Psychological State” and Methodological Solipsism

In order to show this, we need first to clarify the traditional notion of a psychological state. In one sense a state is simply a two-place predicate whose arguments are an individual and a time. In this sense, *being 5 feet tall*, *being in pain*,

knowing the alphabet, and even being a thousand miles from Paris are all states. (Note that the *time* is usually left implicit or "contextual"; the full form of an atomic sentence of these predicates would be "x is five feet tall at time t," "x is in pain at time t," etc.) In science, however, it is customary to restrict the term state to properties which are defined in terms of the parameters of the individual which are fundamental from the point of view of the given science. Thus, being five feet tall is a state (from the point of view of physics); being in pain is a state (from the point of view of mentalistic psychology, at least); knowing the alphabet might be a state (from the point of view of cognitive psychology), although it is hard to say; but being a thousand miles from Paris would *not* naturally be called a *state*. In one sense, a psychological state is simply a state which is studied or described by psychology. In this sense it may be trivially true that, say *knowing the meaning of the word "water"* is a "psychological state" (viewed from the standpoint of cognitive psychology). But this is not the sense of psychological state that is at issue in the above assumption (I).

When traditional philosophers talked about psychological states (or "mental" states), they made an assumption which we may call the assumption of methodological solipsism. This assumption is the assumption that no psychological state, properly so called, presupposes the existence of any individual other than the subject to whom that state is ascribed. (In fact, the assumption was that no psychological state presupposes the existence of the subject's *body* even: if *P* is a psychological state, properly so called, then it must be logically possible for a "disembodied mind" to be in *P*.) This assumption is pretty explicit in Descartes, but it is implicit in just about the whole of traditional philosophical psychology. Making this assumption is, of course, adopting a *restrictive program*—a program which deliberately limits the scope and nature of psychology to fit certain mentalistic preconceptions or, in some cases, to fit an idealistic reconstruction of knowledge and the world. Just *how* restrictive the program is, however, often goes unnoticed. Such common or garden variety psychological states as *being jealous* have to be reconstructed, for example, if the assumption of methodological solipsism is retained. For, in its ordinary use, *x is jealous of y* entails that *y* exists, and *x is jealous of y's regard for z* entails that both *y* and *z* exist (as well

as *x*, of course). Thus *being jealous* and *being jealous of someone's regard for someone else* are not psychological states permitted by the assumption of methodological solipsism. (We shall call them "psychological states in the wide sense" and refer to the states which are permitted by methodological solipsism as "psychological states in the narrow sense.") The reconstruction required by methodological solipsism would be to reconstrue *jealousy* so that I can be jealous of my own hallucinations, or of figments of my imagination, etc. Only if we assume that psychological states in the narrow sense have a significant degree of causal closure (so that restricting ourselves to psychological states in the narrow sense will facilitate the statement of psychological laws) is there any point in engaging in this reconstruction, or in making the assumption of methodological solipsism. But the three centuries of failure of mentalistic psychology is tremendous evidence against this procedure, in my opinion.

Be that as it may, we can now state more precisely what we claimed at the end of the preceding section. Let *A* and *B* be any two terms which differ in extension. By assumption (II) they must differ in meaning (in the sense of "intension"). By assumption (I), *knowing the meaning of A* and *knowing the meaning of B* are psychological states *in the narrow sense*—for this is how we shall construe assumption (I). *But these psychological states must determine the extension of the terms A and B just as much as the meanings ("intensions") do.*

To see this, let us try assuming the opposite. Of course, there cannot be two terms *A* and *B* such that *knowing the meaning of A* is the same state as *knowing the meaning of B* even though *A* and *B* have different extensions. For *knowing the meaning of A* isn't just "grasping the intension" of *A*, whatever that may come to; it is also knowing that the "intension" that one has "grasped" is the intension of *A*. Thus, someone who knows the meaning of "wheel" presumably "grasps the intension" of its German synonym *Rad*; but if he doesn't know that the "intension" in question is the intension of *Rad* he isn't said to "know the meaning of *Rad*." If *A* and *B* are different terms, then *knowing the meaning of A* is a different state from *knowing the meaning of B* whether the meanings of *A* and *B* be themselves the same or different. But by the same argument, if *I*₁ and *I*₂ are different *intensions* and *A* is a term, then *knowing that I*₁ *is the meaning of A* is a different psychological state from

knowing that I_2 is the meaning of A . Thus, there cannot be two different logically possible worlds L_1 and L_2 such that, say, Oscar is in the same psychological state (in the narrow sense) in L_1 and in L_2 (in all respects), but in L_1 Oscar understands A as having the meaning I_1 and in L_2 Oscar understands A as having the meaning I_2 . (For, if there were, then in L_1 Oscar would be in the psychological state *knowing that I_1 is the meaning of A* and in L_2 Oscar would be in the psychological state *knowing that I_2 is the meaning of A* , and these are different and even—assuming that A has just one meaning for Oscar in each world—incompatible psychological states in the narrow sense.)

In short, if S is the sort of psychological state we have been discussing—a psychological state of the form *knowing that I is the meaning of A* , where I is an “intension” and A is a term—then the same necessary and sufficient condition for falling into the extension of A “works” in every logically possible world in which the speaker is in the psychological state S . For the state S determines the intension I , and by assumption (II) the intension amounts to a necessary and sufficient condition for membership in the extension.

If our interpretation of the traditional doctrine of intension and extension is fair to Frege and Carnap, then the whole psychologism/Platonism issue appears somewhat a tempest in a teapot, as far as meaning-theory is concerned. (Of course, it is a very important issue as far as general philosophy of mathematics is concerned.) For even if meanings are “Platonic” entities rather than “mental” entities on the Frege–Carnap view, “grasping” those entities is presumably a psychological state (in the narrow sense). Moreover, the psychological state uniquely determines the “Platonic” entity. So whether one takes the “Platonic” entity or the psychological state as the “meaning” would appear to be somewhat a matter of convention. And taking the psychological state to be the meaning would hardly have the consequence that Frege feared, that meanings would cease to be public. For psychological states are “public” in the sense that different people (and even people in different epochs) can be in the same psychological state. Indeed, Frege’s argument against psychologism is only an argument against identifying concepts with mental particulars, not with mental entities in general.

The “public” character of psychological states entails, in particular, that if Oscar and Elmer understand a word A differently, then they must be

in different psychological states. For the state of *knowing the intension of A to be, say, I* is the same state whether Oscar or Elmer be in it. Thus two speakers cannot be in the same psychological state in all respects and understand the term A differently; the psychological state of the speaker determines the intension (and hence, by assumption (II), the extension) of A .

It is this last consequence of the joint assumptions (I), (II) that we claim to be false. We claim that it is possible for two speakers to be in exactly the same psychological state (in the narrow sense), even though the extension of the term A in the idiolect of the one is different from the extension of the term A in the idiolect of the other. Extension is not determined by psychological state.

This will be shown in detail in later sections. If this is right, then there are two courses open to one who wants to rescue at least one of the traditional assumptions; to give up the idea that psychological state (in the narrow sense) determines intension, or to give up the idea that intension determines extension. We shall consider these alternatives later.

Are Meanings in the Head?

That psychological state does not determine extension will now be shown with the aid of a little science-fiction. For the purpose of the following science-fiction examples, we shall suppose that somewhere in the galaxy there is a planet we shall call Twin Earth. Twin Earth is very much like Earth; in fact, people on Twin Earth even speak *English*. In fact, apart from the differences we shall specify in our science-fiction examples, the reader may suppose that Twin Earth is *exactly* like Earth. He may even suppose that he has a *Doppelgänger*—an identical copy—on Twin Earth, if he wishes, although my stories will not depend on this.

Although some of the people on Twin Earth (say, the ones who call themselves “Americans” and the ones who call themselves “Canadians” and the ones who call themselves “Englishmen,” etc.) speak English, there are, not surprisingly, a few tiny differences which we will now describe between the dialects of English spoken on Twin Earth and Standard English. These differences themselves depend on some of the peculiarities of Twin Earth.

One of the peculiarities of Twin Earth is that the liquid called “water” is not H_2O but a differ-

ent liquid whose chemical formula is very long and complicated. I shall abbreviate this chemical formula simply as XYZ. I shall suppose that XYZ is indistinguishable from water at normal temperatures and pressures. In particular, it tastes like water and it quenches thirst like water. Also, I shall suppose that the oceans and lakes and seas of Twin Earth contain XYZ and not water, that it rains XYZ on Twin Earth and not water, etc.

If a spaceship from Earth ever visits Twin Earth, then the supposition at first will be that "water" has the same meaning on Earth and on Twin Earth. This supposition will be corrected when it is discovered that "water" on Twin Earth is XYZ, and the Earthian spaceship will report somewhat as follows:

"On Twin Earth the word 'water' means XYZ."

(It is this sort of use of the word "means" which accounts for the doctrine that extension is one sense of "meaning," by the way. But note that although "means" does mean something like *has as extension* in this example, one would *not* say

"On Twin Earth the meaning of the word 'water' is XYZ."

unless, possibly, the fact that "water is XYZ" was known to every adult speaker of English on Twin Earth. We can account for this in terms of the theory of meaning we develop below; for the moment we just remark that although the verb "means" sometimes means "has as extension," the nominalization "meaning" *never* means "extension.")

Symmetrically, if a spaceship from Twin Earth ever visits Earth, then the supposition at first will be that the word "water" has the same meaning on Twin Earth and on Earth. This supposition will be corrected when it is discovered that "water" on Earth is H₂O, and the Twin Earthian spaceship will report

"On Earth³ the word 'water' means H₂O."

Note that there is no problem about the extension of the term "water." The word simply has two different meanings (as we say) in the sense in which it is used on Twin Earth, the sense of water_{TE}, what *we* call "water" simply isn't water; while in the sense in which it is used on Earth, the sense of water_E, what the Twin Earthians call "water" simply isn't water. The extension of "water" in the sense of water_E is the set

of all wholes consisting of H₂O molecules, or something like that; the extension of water in the sense of water_{TE} is the set of all wholes consisting of XYZ molecules, or something like that.

Now let us roll the time back to about 1750. At that time chemistry was not developed on either Earth or Twin Earth. The typical Earthian speaker of English did not know water consisted of hydrogen and oxygen, and the typical Twin Earthian speaker of English did not know "water" consisted of XYZ. Let Oscar₁ be such a typical Earthian English speaker, and let Oscar₂ be his counterpart on Twin Earth. You may suppose that there is no belief that Oscar₁ had about water that Oscar₂ did not have about "water." If you like, you may even suppose that Oscar₁ and Oscar₂ were exact duplicates in appearance, feelings, thoughts, interior monologue, etc. Yet the extension of the term "water" was just as much H₂O on Earth in 1750 as in 1950; and the extension of the term "water" was just as much XYZ on Twin Earth in 1750 as in 1950. Oscar₁ and Oscar₂ understood the term "water" differently in 1750 *although they were in the same psychological state*, and although, given the state of science at the time, it would have taken their scientific communities about fifty years to discover that they understood the term "water" differently. Thus the extension of the term "water" (and, in fact, its "meaning" in the intuitive preanalytical usage of that term) is *not* a function of the psychological state of the speaker by itself.

But, it might be objected, why should we accept it that the term "water" has the same extension in 1750 and in 1950 (on both Earths)? The logic of natural-kind terms like "water" is a complicated matter, but the following is a sketch of an answer. Suppose I point to a glass of water and say "this liquid is called water" (or "this is called water," if the marker "liquid" is clear from the context). My "ostensive definition" of water has the following empirical presupposition that the body of liquid I am pointing to bears a certain sameness relation (say, *x is the same liquid as y*, or *x is the same_L as y*) to most of the stuff I and other speakers in my linguistic community have on other occasions called "water." If this presupposition is false because, say, I am without knowing it pointing to a glass of gin and not a glass of water, then I do not intend my ostensive definition to be accepted. Thus the ostensive definition conveys what might be called a defeasible necessary and suf-

ficient condition: the necessary and sufficient condition for being water is bearing the relation same_L to the stuff in the glass; but this is the necessary and sufficient condition only if the empirical presupposition is satisfied. If it is not satisfied, then one of a series of, so to speak, “fallback” conditions becomes activated.

The key point is that the relation same_L is a *theoretical* relation whether something is or is not the same liquid as *this* may take an indeterminate amount of scientific investigation to determine. Moreover, even if a “definite” answer has been obtained either through scientific investigation or through the application of some “common sense” test, the answer is *defeasible*: future investigation might reverse even the most “certain” example. Thus, the fact that an English speaker in 1750 might have called XYZ “water,” while he or his successors would not have called XYZ water in 1800 or 1850 does not mean that the “meaning” of “water” changed for the average speaker in the interval. In 1750 or in 1850 or in 1950 one might have pointed to, say, the liquid in Lake Michigan as an example of “water.” What changed was that in 1750 we would have mistakenly thought that XYZ bore the relation same_L to the liquid in Lake Michigan, while in 1800 or 1850 we would have known that it did not (I am ignoring the fact that the liquid in Lake Michigan was only dubiously water in 1950, of course).

Let us now modify our science-fiction story. I do not know whether one can make pots and pans out of molybdenum; and if one can make them out of molybdenum, I don’t know whether they could be distinguished easily from aluminum pots and pans (I don’t know any of this even though I have acquired the word “molybdenum.”) So I shall suppose that molybdenum pots and pans *can’t* be distinguished from aluminum pots and pans save by an expert. (To emphasize the point, I repeat that this could be true for all I know, and *a fortiori* it could be true for all I know by virtue of “knowing the meaning” of the words *aluminum* and *molybdenum*.) We will now suppose that molybdenum is as common on Twin Earth as aluminum is on Earth, and that aluminum is as rare on Twin Earth as molybdenum is on Earth. In particular, we shall assume that “aluminum” pots and pans are made of molybdenum on Twin Earth. Finally, we shall assume that the words “aluminum” and “molybdenum” are *switched* on Twin Earth: “aluminum” is the name of *molybdenum* and “molybdenum” is the name of *aluminum*.

This example shares some features with the

previous one. If a spaceship from Earth visited Twin Earth, the visitors from Earth probably would not suspect that the “aluminum” pots and pans on Twin Earth were not made of aluminum, especially when the Twin Earthians *said* they were. But there is one important difference between the two cases. An Earthian metallurgist could tell very easily that “aluminum” was molybdenum, and a Twin Earthian metallurgist could tell equally easily that aluminum was “molybdenum.” (The shudder quotes in the preceding sentence indicate Twin Earthian usages.) Whereas in 1750 no one on either Earth or Twin Earth could have distinguished water from “water,” the confusion of aluminum with “aluminum” involves only a part of the linguistic communities involved.

The example makes the same point as the preceding one. If Oscar₁ and Oscar₂ are standard speakers of Earthian English and Twin Earthian English respectively, and neither is chemically or metallurgically sophisticated, then there may be no difference at all in their psychological state when they use the word “aluminum”; nevertheless we have to say that “aluminum” has the extension *aluminum* in the idiolect of Oscar₁ and the extension *molybdenum* in the idiolect of Oscar₂. (Also we have to say that Oscar₁ and Oscar₂ mean different things by “aluminum,” that “aluminum” has a different meaning on Earth than it does on Twin Earth, etc.) Again we see that the psychological state of the speaker does *not* determine the extension (*or* the “meaning,” speaking preanalytically) of the word.

Before discussing this example further, let me introduce a *non-science-fiction* example. Suppose you are like me and cannot tell an elm from a beech tree. We still say that the extension of “elm” in my idiolect is the same as the extension of “elm” in anyone else’s, viz., the set of all elm trees, and that the set of all beech trees is the extension of “beech” in *both* of our idiolects. Thus “elm” in my idiolect has a different extension from “beech” in your idiolect (as it should). Is it really credible that this difference in extension is brought about by some difference in our *concepts*? My concept of an elm tree is exactly the same as my concept of a beech tree (I blush to confess). (This shows that the identification of meaning “in the sense of intension” with *concept* cannot be correct, by the way.) If someone heroically attempts to maintain that the difference between the extension of “elm” and the extension of “beech” in my idiolect is explained by a difference in my psychological state, then we can always refute him by constructing a “Twin

Earth" example—just let the words "elm" and "beech" be switched on Twin Earth (the way "aluminum" and "molybdenum" were in the previous example). Moreover, I suppose I have a *Doppelgänger* on Twin Earth who is molecule for molecule "identical" with me (in the sense in which two neckties can be "identical"). If you are a dualist, then also suppose my *Doppelgänger* thinks the same verbalized thoughts I do, has the same sense data, the same dispositions, etc. It is absurd to think *his* psychological state is one bit different from mine: yet he "means" *beech* when he says "elm" and *I* "mean" *elm* when I say elm. Cut the pie any way you like, "meanings" just ain't in the *head!*

A Sociolinguistic Hypothesis

The last two examples depend upon a fact about language that seems, surprisingly, never to have been pointed out: that there is *division of linguistic labor*. We could hardly use such words as "elm" and "aluminum" if no one possessed a way of recognizing elm trees and aluminum metal; but not everyone to whom the distinction is important has to be able to make the distinction. Let us shift the example: consider *gold*. Gold is important for many reasons: it is a precious metal, it is a monetary metal, it has symbolic value (it is important to most people that the "gold" wedding ring they wear *really* consist of gold and not just *look* gold), etc. Consider our community as a "factory": in this "factory" some people have the "job" of *wearing gold wedding rings*, other people have the "job" of *selling gold wedding rings*, still other people have the "job" of *telling whether or not something is really gold*. It is not at all necessary or efficient that everyone who wears a gold ring (or a gold cufflink, etc.), or discusses the "gold standard," etc., engage in buying and selling gold. Nor is it necessary or efficient that everyone who buys and sells gold be able to tell whether or not something is really gold in a society where this form of dishonesty is uncommon (selling fake gold) and in which one can easily consult an expert in case of doubt. And it is *certainly* not necessary or efficient that everyone who has occasion to buy or wear gold be able to tell with any reliability whether or not something is really gold.

The foregoing facts are just examples of mundane division of labor (in a wide sense). But they engender a division of linguistic labor: everyone to whom gold is important for any rea-

son has to *acquire* the word "gold"; but he does not have to acquire the *method of recognizing* if something is or is not gold. He can rely on a special subclass of speakers. The features that are generally thought to be present in connection with a general name—necessary and sufficient conditions for membership in the extension, ways of recognizing if something is in the extension ("criteria"), etc.—are all present in the linguistic community *considered as a collective body*; but that collective body divides the "labor" of knowing and employing these various parts of the "meaning" of "gold."

This division of linguistic labor rests upon and presupposes the division of *nonlinguistic* labor, of course. If only the people who know how to tell if some metal is really gold or not have any reason to have the word "gold" in their vocabulary, then the word "gold" will be as the word "water" was in 1750 with respect to that subclass of speakers, and the other speakers just won't acquire it at all. And some words do not exhibit any division of linguistic labor: "chair," for example. But with the increase of division of labor in the society and the rise of science, more and more words begin to exhibit this kind of division of labor. "Water," for example, did not exhibit it at all prior to the rise of chemistry. Today it is obviously necessary for every speaker to be able to recognize water (reliably under normal conditions), and probably every adult speaker even knows the necessary and sufficient condition "water is H₂O," but only a few adult speakers could distinguish water from liquids which superficially resembled water. In case of doubt, other speakers would rely on the judgement of these "expert" speakers. Thus the way of recognizing possessed by these "expert" speakers is also, through them, possessed by the collective linguistic body, even though it is not possessed by each individual member of the body, and in this way the most *recherché* fact about water may become part of the *social* meaning of the word while being unknown to almost all speakers who acquire the word.

It seems to me that this phenomenon of division of linguistic labor is one which it will be very important for sociolinguistics to investigate. In connection with it, I should like to propose the following hypothesis:

HYPOTHESIS OF THE UNIVERSALITY OF THE DIVISION OF LINGUISTIC LABOR: Every linguistic community exemplifies the sort of division of linguistic labor just described: that is, possesses at least some terms whose associated "criteria" are known only to a subset of

the speakers who acquire the terms, and whose use by the other speakers depends upon a structured cooperation between them and the speakers in the relevant subsets.

It would be of interest, in particular, to discover if extremely primitive peoples were sometimes exceptions to this hypothesis (which would indicate that the division of linguistic labor is a product of social evolution), or if even they exhibit it. In the latter case, one might conjecture that division of labor, including linguistic labor, is a fundamental trait of our species.

It is easy to see how this phenomenon accounts for some of the examples given above of the failure of the assumptions (I), (II). Whenever a term is subject to the division of linguistic labor, the “average” speaker who acquires it does not acquire anything that fixes its extension. In particular, his individual psychological state *certainly* does not fix its extension; it is only the sociolinguistic state of the collective linguistic body to which the speaker belongs that fixes the extension.

We may summarize this discussion by pointing out that there are two sorts of tools in the world: there are tools like a hammer or a screwdriver which can be used by one person; and there are tools like a steamship which require the cooperative activity of a number of persons to use. Words have been thought of too much on the model of the first sort of tool.

Indexicality and Rigidity⁴

The first of our science-fiction examples—“water” on Earth and on Twin Earth in 1750—does not involve division of linguistic labor, or at least does not involve it in the same way the examples of “aluminum” and “elm” do. There were not (in our story, anyway) any “experts” on water on Earth in 1750, nor any experts on “water” on Twin Earth. (The example *can* be construed as involving division of labor *across time*, however. I shall not develop this method of treating the example here.) The example *does* involve things which are of fundamental importance to the theory of reference and also to the theory of necessary truth, which we shall now discuss.

There are two obvious ways of telling someone what one means by a natural-kind term such as “water” or “tiger” or “lemon.” One can give him a so-called ostensive definition—“this (liquid) is water”; “this (animal) is a tiger”; “this

(fruit) is a lemon”; where the parentheses are meant to indicate that the “markers” *liquid, animal, fruit*, may be either explicit or implicit. Or one can give him a *description*. In the latter case the description one gives typically consists of one or more markers together with a *stereotype*⁵—a standardized description of features of the kind that are typical, or “normal,” or at any rate stereotypical. The central features of the stereotype generally are *criteria*—features which in normal situations constitute ways of recognizing if a thing belongs to the kind or, at least, necessary conditions (or probabilistic necessary conditions) for membership in the kind. Not all criteria used by the linguistic community as a collective body are included in the stereotype, and in some cases the stereotypes may be quite weak. Thus (unless I am a very atypical speaker), the stereotype of an elm is just that of a common deciduous tree. These features are indeed necessary conditions for membership in the kind (I mean “necessary” in a loose sense; I don’t think “elm trees are deciduous” is *analytic*), but they fall far short of constituting a way of recognizing elms. On the other hand, the stereotype of a tiger does enable one to recognize tigers (unless they are albino, or some other atypical circumstance is present), and the stereotype of a lemon generally enables one to recognize lemons. In the extreme case, the stereotype may be *just* the marker: the stereotype of molybdenum might be *just* that molybdenum is a *metal*. Let us consider both of these ways of introducing a term into someone’s vocabulary.

Suppose I point to a glass of liquid and say “*this* is water,” in order to teach someone the word “water.” We have already described some of the empirical presuppositions of this act, and the way in which this kind of meaning-explanation is defeasible. Let us now try to clarify further how it is supposed to be taken.

In what follows, we shall take the notion of “possible world” as primitive. We do this because we feel that in several senses the notion makes sense and is scientifically important even if it needs to be made more precise. We shall assume further that in at least some cases it is possible to speak of the same individual as existing in more than one possible world.⁶ Our discussion leans heavily on the work of Saul Kripke, although the conclusions were obtained independently.

Let W_1 and W_2 be two possible worlds in which I exist and in which this glass exists and in which I am giving a meaning explanation by

pointing to this glass and saying "this is water." (We do *not* assume that the *liquid* in the glass is the same in both worlds.) Let us suppose that in W_1 the glass is full of H_2O and in W_2 the glass is full of XYZ . We shall also suppose that W_1 is the actual world and that XYZ is the stuff typically called "water" in the world W_2 (so that the relation between English speakers in W_1 and English speakers in W_2 is exactly the same as the relation between English speakers on Earth and English speakers on Twin Earth). Then there are two theories one might have concerning the meaning of "water":

1. One might hold that "water" was *world-relative* but *constant* in meaning (i.e., the word has a *constant relative meaning*). In this theory, "water" means the same in W_1 and W_2 ; it's just that water is H_2O in W_1 and water is XYZ in W_2 .
2. One might hold that water is H_2O in all worlds (the stuff called "water" in W_2 isn't water), but "water" doesn't have the same meaning in W_1 and W_2 .

If what was said before about the Twin Earth case was correct, then (2) is clearly the correct theory. When I say "this (liquid) is water," the "this" is, so to speak, a *de re* "this"—i.e., the force of my explanation is that "water" is whatever bears a certain equivalence relation (the relation we called "same_L" above) to the piece of liquid referred to as "this" in the actual world.

We might symbolize the difference between the two theories as a "scope" difference in the following way. In theory (1), the following is true:

- (1') (For every world W) (For every x in W)
 (x is water $\equiv x$ bears same_L to the entity referred to as "this" in W)

while on theory (2):

- (2') (For every world W) (For every x in W)
 (x is water $\equiv x$ bears same_L to the entity referred to as "this" in the actual world W_1).

(I call this a "scope" difference because in (1') "the entity referred to as 'this'" is within the scope of "For every world W "—as the qualifying phrase "in W " makes explicit, whereas in (2') "the entity referred to as 'this'" means "the entity referred to as 'this' in the actual world," and has thus a reference *independent* of the bound variable " W .")

Kripke calls a designator "rigid" (in a given sentence) if (in that sentence) it refers to the same individual in every possible world in which the designator designates. If we extend the notion of rigidity to substance names, then we may express Kripke's theory and mine by saying that the term "water" is *rigid*.

The rigidity of the term "water" follows from the fact that when I give the ostensive definition "this (liquid) is water" I intend (2') and not (1').

We may also say, following Kripke, that when I give the ostensive definition "this (liquid) is water," the demonstrative "this" is *rigid*.

What Kripke was the first to observe is that this theory of the meaning (or "use," or whatever) of the word "water" (and other natural-kind terms as well) has startling consequences for the theory of necessary truth.

To explain this, let me introduce the notion of a *cross-world relation*. A two term relation R will be called *cross-world* when it is understood in such a way that its extension is a set of ordered pairs of individuals *not all in the same possible world*. For example, it is easy to understand the relation *same height as* as a cross-world relation: just understand it so that, e.g., if x is an individual in a world W_1 who is five feet tall (in W_1) and y is an individual in W_2 who is five feet tall (in W_2), then the ordered pair x, y belongs to the extension of *same height as*. (Since an individual may have different heights in different possible worlds in which that same individual exists, strictly speaking it is not the ordered pair x, y that constitutes an element of the extension of *same height as*, but rather the ordered pair x -in-world- W_1, y -in-world- W_2 .)

Similarly, we can understand the relation *same_L* (same liquid as) as a cross-world relation by understanding it so that a liquid in world W_1 which has the same important physical properties (in W_1) that a liquid in W_2 possesses (in W_2) bears *same_L* to the latter liquid.

Then the theory we have been presenting may be summarized by saying that an entity x , in an arbitrary possible world, is *water* if and only if it bears the relation *same_L* (construed as a cross-world relation) to the stuff we call "water" in the actual world.

Suppose, now, that I have not yet discovered what the important physical properties of water are (in the actual world)—i.e., I don't yet know that water is H_2O . I may have ways of *recognizing* water that are successful (of course, I may make a small number of mistakes that I won't be able to detect until a later stage in our scientific

development) but not know the microstructure of water. If I agree that a liquid with the superficial properties of “water” but a different microstructure *isn't really water*, then my ways of recognizing water (my “operational definition,” so to speak) cannot be regarded as an analytical specification of *what it is to be* water. Rather, the operational definition, like the ostensive one, is simply a way of pointing out a standard—pointing out the stuff *in the actual world* such that for x to be water, in *any* world, is for x to bear the relation same_L to the *normal* members of the class of *local* entities that satisfy the operational definition. “Water” on Twin Earth is not water, even if it satisfies the operational definition, because it doesn't bear same_L to the *local* stuff that satisfies the operational definition, and local stuff that satisfies the operational definition but has a microstructure different from the rest of the local stuff that satisfies the operational definition isn't water either, because it doesn't bear same_L to the *normal* examples of the local “water.”

Suppose, now, that I discover the microstructure of water—that water is H_2O . At this point I will be able to say that the stuff on Twin Earth that I earlier *mistook* for water isn't really water. In the same way if you describe not another planet in the actual universe, but another possible universe in which there is stuff with the chemical formula XYZ which passes the “operational test” for *water*, we shall have to say that that stuff isn't water but merely XYZ . You will not have described a possible world in which “water is XYZ ,” but merely a possible world in which there are lakes of XYZ , people drink XYZ (and not water), or whatever. In fact, once we have discovered the nature of water, nothing counts as a possible world in which water doesn't have that nature. Once we have discovered that water (in the actual world) is H_2O , *nothing counts as a possible world in which water isn't H_2O* . In particular, if a “logically possible” statement is one that holds in some “logically possible world,” *it isn't logically possible that water isn't H_2O* .

On the other hand, we can perfectly well imagine having experiences that would convince us (and that would make it rational to believe that) water *isn't* H_2O . In that sense, it is conceivable that water isn't H_2O . It is conceivable but it isn't logically possible! Conceivability is no proof of logical possibility.

Kripke refers to statements which are rationally unrevisable (assuming there are such) as

epistemically necessary. Statements which are true in all possible worlds he refers to simply as necessary (or sometimes as “metaphysically necessary”). In this terminology, the point just made can be restated as: a statement can be (metaphysically) necessary and epistemically contingent. Human intuition has no privileged access to metaphysical necessity.

Since Kant there has been a big split between philosophers who thought that all necessary truths were analytic and philosophers who thought that some necessary truths were synthetic a priori. But none of these philosophers thought that a (metaphysically) necessary truth could fail to be a priori: the Kantian tradition was as guilty as the empiricist tradition of equating metaphysical and epistemic necessity. In this sense Kripke's challenge to received doctrine goes far beyond the usual empiricism/Kantianism oscillation.

In this paper our interest is in theory of meaning, however, and not in theory of necessary truth. Points closely related to Kripke's have been made in terms of the notion of *indexicality*.⁷ Words like “now,” “this,” “here,” have long been recognized to be *indexical*, or *token-reflexive*—i.e., to have an extension which varied from context to context or token to token. For these words no one has ever suggested the traditional theory that “intension determines extension.” To take our Twin Earth example: if I have a *Doppelgänger* on Twin Earth, then when I think “I have a headache,” *he* thinks “I have a headache.” But the extension of the particular token of “I” in his verbalized thought is himself (or his unit class, to be precise), while the extension of the token of “I” in *my* verbalized thought is *me* (or my unit class, to be precise). So the same word, “I,” has two different extensions in two different idiolects; but it does not follow that the concept I have of myself is in any way different from the concept my *Doppelgänger* has of himself.

Now then, we have maintained that indexicality extends beyond the *obviously* indexical words and morphemes (e.g., the tenses of verbs). Our theory can be summarized as saying that words like “water” have an unnoticed indexical component: “water” is stuff that bears a certain similarity relation to the water *around here*. Water at another time or in another place or even in another possible world has to bear the relation same_L to *our* “water” *in order to be water*. Thus the theory that (1) words have “intensions,” which are something like concepts

associated with the words by speakers; and that (2) intension determines extension—cannot be true of natural-kind words like “water” for the same reason the theory cannot be true of obviously indexical words like “I.”

The theory that natural-kind words like “water” are indexical leaves it open, however, whether to say that “water” in the Twin Earth dialect of English has the same *meaning* as “water” in the Earth dialect and a different extension (which is what we normally say about “I” in different idiolects), thereby giving up the doctrine that “meaning (intension) determines extension”; or to say, as we have chosen to do, that difference in extension is *ipso facto* a difference in meaning for natural-kind words, thereby giving up the doctrine that meanings are concepts, or, indeed, mental entities of *any* kind.

It should be clear, however, that Kripke’s doctrine that natural-kind words are rigid designators and our doctrine that they are indexical are but two ways of making the same point. We heartily endorse what Kripke says when he writes:

Let us suppose that we do fix the reference of a name by a description. Even if we do so, we do not then make the name synonymous with the description, but instead we use the name rigidly to refer to the object so named, even in talking about counterfactual situations where the thing named would not satisfy the description in question. Now, this is what I think is in fact true for those cases of naming where the reference is fixed by description. But, in fact, I also think, contrary to most recent theorists, that the reference of names is rarely or almost never fixed by means of description. And by this I do not just mean what Searle says: “It’s not a single description, but rather a cluster, a family of properties that fixes the reference.” I mean that properties in this sense are not used at all.⁸

Other Words

... So far we have only used natural-kind words as examples, but the points we have made apply to many other kinds of words as well. They apply to the great majority of all nouns, and to other parts of speech as well.

Let us consider for a moment the names of artifacts—words like “pencil,” “chair,” “bottle,” etc. The traditional view is that these words are certainly defined by conjunctions, or possibly clusters, of properties. Anything with all of the properties in the conjunction (or sufficiently

many of the properties in the cluster, on the cluster model) is necessarily a *pencil*, *chair*, *bottle*, or whatever. In addition, some of the properties in the cluster (on the cluster model) are usually held to be *necessary* (on the conjunction-of-properties model, *all* of the properties in the conjunction are necessary). *Being an artifact* is supposedly necessary, and belonging to a kind with a certain standard purpose—e.g., “pencils are artifacts,” and “pencils are standardly intended to be written with” are supposed to be necessary. Finally, this sort of necessity is held to be *epistemic* necessity—in fact, analyticity.

Let us once again engage in science fiction. This time we use an example devised by Rogers Albritton. Imagine that we someday discover that *pencils are organisms*. We cut them open and examine them under the electron microscope, and we see the almost invisible tracery of nerves and other organs. We spy upon them, and we see them spawn, and we see the offspring grow into full-grown pencils. We discover that these organisms are not imitating other (artificial) pencils—there are not and never were any pencils except these organisms. It is strange, to be sure, that there is *lettering* on many of these organisms—e.g., BONDED Grants DELUXE made in U.S.A. No. 2.—perhaps they are intelligent organisms, and this is their form of camouflage. (We also have to explain why no one ever attempted to manufacture pencils, etc., but this is clearly a possible world, in some sense.)

If this is conceivable, and I agree with Albritton that it is, then it is epistemically possible that *pencils could turn out to be organisms*. It follows that *pencils are artifacts* is not epistemically necessary in the strongest sense and, a fortiori, not analytic.

Let us be careful, however. Have we shown that there is a possible world in which pencils are organisms? I think not. What we have shown is that there is a possible world in which certain organisms are the *epistemic counterparts* of pencils (the phrase is Kripke’s). To return to the device of Twin Earth: imagine this time that pencils on Earth are just what we think they are, artifacts manufactured to be written with, while “pencils” on Twin Earth are organisms à la Albritton. Imagine, further, that this is totally unsuspected by the Twin Earthians—they have exactly the beliefs about “pencils” that we have about pencils. When we discovered this, we would not say: “some pencils are organisms.” We would be far more likely to say: “the things

on Twin Earth that pass for pencils aren't really pencils. They're really a species of organism."

Suppose now the situation to be as in Albritton's example both on Earth and on Twin Earth. Then we would say "pencils are organisms." Thus, whether the "pencil-organisms" on Twin Earth (or in another possible universe) are really *pencils* or not is a function of whether or not the *local* pencils are organisms or not. If the local pencils are just what we think they are, then a possible world in which there are pencil-organisms is *not* a possible world in which *pencils are organisms*; there are *no* possible worlds in which pencils are organisms in this case (which is, of course, the actual one). That pencils are artifacts is necessary in the sense of true in all possible worlds—metaphysically necessary. But it doesn't follow that it's epistemically necessary.

It follows that "pencil" is not *synonymous* with any description—not even loosely synonymous with a *loose* description. When we use the word "pencil," we intend to refer to whatever has the same *nature* as the normal examples of the local pencils in the actual world. "Pencil" is just as *indexical* as "water" or "gold."

In a way, the case of pencils turning out to be organisms is complementary to the case we discussed some years ago⁹ of cats turning out to be robots (remotely controlled from Mars). Katz¹⁰ argues that we misdescribed this case: that the case should rather be described as its *turning out that there are no cats in this world*. Katz admits that we might say "Cats have turned out not to be animals, but robots"; but he argues that this is a semantically deviant sentence which is glossed as "the things I am referring to as 'cats' have turned out not to be animals, but robots." Katz's theory is bad linguistics, however. First of all, the explanation of how it is we can say "Cats are robots" is simply an all-purpose explanation of how we can say *anything*. More important, Katz's theory predicts that "Cats are robots" is *deviant*, while "There are no cats in the world" is nondeviant, in fact standard, in the case described. Now then, I don't deny that there *is* a case in which "There are not (and never were) any cats in the world" would be standard: we might (speaking epistemically) discover that we have been suffering from a collective hallucination. ("Cats" are like pink elephants.) But in the case I described, "Cats have turned out to be robots remotely controlled from Mars" is surely nondeviant, and "There are no cats in the world" is highly deviant.

Incidentally, Katz's account is not only bad linguistics; it is also bad as a rational reconstruction. The reason we *don't* use "cat" as synonymous with a description is surely that we know enough about cats to know that they do have a hidden structure, and it is good scientific methodology to use the name to refer rigidly to the things that possess that hidden structure, and not to whatever happens to satisfy some description. Of course, if we *knew* the hidden structure we could frame a description in terms of *it*; but we don't at this point. In this sense the use of natural-kind words reflects an important fact about our relation to the world: we know that there are kinds of things with common hidden structure, but we don't yet have the knowledge to describe all those hidden structures.

Katz's view has more plausibility in the "pencil" case than in the "cat" case, however. We think we *know* a necessary and sufficient condition for being a *pencil*, albeit a vague one. So it is possible to make "pencil" synonymous with a loose description. We *might* say, in the case that "pencils turned out to be organisms" *either* "Pencils have turned out to be organisms" *or* "There are no pencils in the world"—i.e., we might use "pencil" either as a natural-kind word or as a "one-criterion" word.¹¹

On the other hand, we might doubt that there *are* any true one-criterion words in natural language, apart from stipulative contexts. Couldn't it turn out that pediatricians aren't doctors but Martian spies? Answer "yes," and you have abandoned the synonymy of "pediatrician" and "doctor specializing in the care of children." It seems that there is a strong tendency for words which are introduced as "one-criterion" words to develop a "natural-kind" sense, with all the concomitant rigidity and indexicality. In the case of artifact-names, this natural-kind sense seems to be the predominant one.

(There is a joke about a patient who is on the verge of being discharged from an insane asylum. The doctors have been questioning him for some time, and he has been giving perfectly sane responses. They decide to let him leave, and at the end of the interview one of the doctors inquires casually, "What do you want to be when you get out?" "A teakettle." The joke would not be intelligible if it were literally inconceivable that a person could be a teakettle.)

There are, however, words which retain an almost pure one-criterion character. These are words whose meaning derives from a transformation: *hunter = one who hunts*.

Not only does the account given here apply to most nouns, but it also applies to other parts of speech. Verbs like "grow," adjectives like "red," etc., all have indexical features. On the other hand, some syncategorematic words seem to have more of a one-criterion character. "Whole," for example, can be explained thus: *The army surrounded the town* could be true even if the A division did not take part. *The whole army surrounded the town* means every part of the army (of the relevant kind, e.g., the A Division) took part in the action signified by the verb.¹²

Meaning

Let us now see where we are with respect to the notion of meaning. We have now seen that the extension of a term is not fixed by a concept that the individual speaker has in his head, and this is true both because extension is, in general, determined *socially*—there is division of linguistic labor as much as of "real" labor—and because extension is, in part, determined *indexically*. The extension of our terms depends upon the actual nature of the particular things that serve as paradigms,¹³ and this actual nature is not, in general, fully known to the speaker. Traditional semantic theory leaves out only two contributions to the determination of extension—the contribution of society and the contribution of the real world!

We saw at the outset that meaning cannot be identified with extension. Yet it cannot be identified with "intension" either, if intension is something like an individual speaker's *concept*. What are we to do?

There are two plausible routes that we might take. One route would be to retain the identification of meaning with concept and pay the price of giving up the idea that meaning determines extension. If we followed this route, we might say that "water" has the same *meaning* on Earth and on Twin Earth, but a different *extension*. (Not just a different *local* extension but a different *global* extension. The XYZ on Twin Earth isn't in the extension of the tokens of "water" that I utter, but it is in the extension of the tokens of "water" that my *Doppelgänger* utters, and this isn't just because Twin Earth is far away from me, since molecules of H₂O are in the extension of the tokens of "water" that I utter no matter how far away from me they are in space and time. Also, what I can counterfactually suppose water to be is different from what

my *Doppelgänger* can counterfactually suppose "water" to be.) While this is the correct route to take for an *absolutely* indexical word like "I," it seems incorrect for the words we have been discussing. Consider "elm" and "beech," for example. If these are "switched" on Twin Earth, then surely we would *not* say that "elm" has the same meaning on Earth and Twin Earth, even if my *Doppelgänger's* stereotype of a beech (or an "elm," as he calls it) is identical with my stereotype of an elm. Rather, we would say that "elm" in my *Doppelgänger's* idiolect means *beech*. For this reason, it seems preferable to take a different route and identify "meaning" with an ordered pair (or possibly an ordered *n-tuple*) of entities, *one of which is the extension*. (The other components of the, so to speak, "meaning vector" will be specified later.) Doing this makes it trivially true that *meaning determines extension* (i.e., difference in extension is ipso facto difference in meaning), but totally abandons the idea that if there is a difference in the meaning my *Doppelgänger* and I assign to a word, then there *must* be some difference in our concepts (or in our psychological state). Following this route, we can say that my *Doppelgänger* and I mean *something different* when we say "elm," but this will not be an assertion about our psychological states. All this means is that the tokens of the word he utters have a different extension than the tokens of the word I utter; but this difference in extension is not a reflection of any difference in our individual linguistic competence considered in isolation.

If this is correct, and I think it is, then the traditional problem of meaning splits into two problems. The first problem is to account for the *determination of extension*. Since, in many cases, extension is determined socially and not individually, owing to the division of linguistic labor, I believe that this problem is properly a problem for sociolinguistics. Solving it would involve spelling out in detail exactly how the division of linguistic labor works. The so-called "causal theory of reference," introduced by Kripke for proper names and extended by us to natural-kind words and physical-magnitude terms, falls into this province. For the fact that, in many contexts, we assign to the tokens of a name that I utter whatever referent we assign to the tokens of the same name uttered by the person from whom I acquired the name (so that the reference is transmitted from speaker to speaker, starting from the speakers who were present at the "naming ceremony," even though no fixed

description is transmitted) is simply a special case of social cooperation in the determination of reference.

The other problem is to describe *individual competence*. Extension may be determined socially, in many cases, but we don't assign the standard extension to the tokens of a word *W* uttered by Jones *no matter how* Jones uses *W*. Jones has to have some particular ideas and skills in connection with *W* in order to play his part in the linguistic division of labor. Once we give up the idea that individual competence has to be so strong as to actually determine extension, we can begin to study it in a fresh frame of mind. . . .

The Meaning of "Meaning"

We may now summarize what has been said in the form of a proposal concerning how one might reconstruct the notion of "meaning." Our proposal is not the only one that might be advanced on the basis of these ideas, but it may serve to encapsulate some of the major points. In addition, I feel that it recovers as much of ordinary usage in common sense talk and in linguistics as one is likely to be able to conveniently preserve. Since, in my view something like the assumptions (I) and (II) listed in the first part of this paper are deeply embedded in ordinary meaning talk, and these assumptions are jointly inconsistent with the facts, no reconstruction is going to be without some counterintuitive consequences.

Briefly, my proposal is to define "meaning" not by picking out an object which will be identified with the meaning (although that might be done in the usual set-theoretic style if one insists), but by specifying a normal form (or, rather, a *type* of normal form) for the description of meaning. If we know what a "normal form description" of the meaning of a word should be, then, as far as I am concerned, we know what meaning is in any scientifically interesting sense.

My proposal is that the normal form description of the meaning of a word should be a finite sequence, or "vector," whose components should certainly include the following (it might be desirable to have other types of components as well): (1) the syntactic markers that apply to the word, e.g., "noun"; (2) the semantic markers that apply to the word, e.g., "animal," "period of time"; (3) a description of the additional fea-

tures of the stereotype, if any; (4) a description of the extension.

The following convention is a part of this proposal: the components of the vector all represent a hypothesis about the individual speaker's competence, *except the extension*. Thus the normal form description for "water" might be, in part:

Syntactic Markers

mass noun; concrete;

Semantic Markers

natural-kind; liquid;

Stereotype

colorless; transparent; tasteless; thirst-quenching; etc.

Extension

H_2O (give or take impurities)

—this does not mean that knowledge of the fact that water is H_2O is being imputed to the individual speaker or even to the society. It means that (*we* say) the extension of the term "water" as *they* (the speakers in question) use it is *in fact* H_2O . The objection "who are *we* to say what the extension of *their* term is in fact" has been discussed above. Note that this is fundamentally an objection to the notion of *truth*, and that extension is a relative of truth and inherits the family problems.

Let us call two descriptions *equivalent* if they are the same except for the description of the extension, and the two descriptions are coextensive. Then, if the set variously described in the two descriptions is, *in fact*, the extension of the word in question, and the other components in the description are correct characterizations of the various aspects of competence they represent, *both* descriptions count as correct. Equivalent descriptions are both correct or both incorrect. This is another way of making the point that, although we have to use a *description* of the extension to *give* the extension, we think of the component in question as being the *extension* (the *set*), not the description of the extension.

In particular the representation of the words "water" in Earth dialect and "water" in Twin Earth dialect would be the same except that in the last column the normal form description of the Twin Earth word "water" would have *XYZ* and not H_2O . This means, in view of what has

just been said, that we are ascribing the *same* linguistic competence to the typical Earthling/Twin Earthian speaker, but a different extension to the word, nonetheless.

This proposal means that we keep assumption (II) of our early discussion. Meaning determines extension—by construction, so to speak. But (I) is given up; the psychological state of the individual speaker does not determine “what he means.”

In most contexts this will agree with the way we speak, I believe. But one paradox: suppose Oscar is a German-English bilingual. In our view, in his total collection of dialects, the words “beechn” and *Buche* are *exact synonyms*. The normal form descriptions of their meanings would be identical. But he might very well not know that they are synonyms! A speaker can have two synonyms in his vocabulary and not know that they are synonyms!

It is instructive to see how the failure of the apparently obvious “if S_1 and S_2 are synonyms and Oscar understands both S_1 and S_2 then Oscar knows that S_1 and S_2 are synonyms” is related to the falsity of (I), in our analysis. Notice that if we had chosen to omit the extension as a component of the “meaning-vector,” which is David Lewis’s proposal as I understand it, then we would have the paradox that “elm” and “beechn” have the *same meaning* but different extensions!

On just about any materialist theory, believing a proposition is likely to involve processing some *representation* of that proposition, be it a sentence in a language, a piece of “brain code,” a thought form, or whatever. Materialists, and not only materialists, are reluctant to think that one can believe propositions *neat*. But even materialists tend to believe that, if one believes a proposition, *which* representation one employs is (pardon the pun) immaterial. If S_1 and S_2 are both representations that are *available* to me, then if I believe the proposition expressed by S_1 under the representation S_1 , I must also believe it under the representation S_2 —at least, I must do this if I have any claim to rationality. But, as we have just seen, this isn’t right. Oscar may well believe that *this* is a “beechn” (it has a sign on it that says “beechn”), but not believe or disbelieve that this is a “*Buche*.” It is not just that belief is a process involving representations; he believes the proposition (if one wants to introduce “propositions” at all) under one representation and not under another.

The amazing thing about the theory of meaning is how long the subject has been in the grip of philosophical misconceptions, and how strong these misconceptions are. Meaning has been identified with a necessary and sufficient condition by philosopher after philosopher. In the empiricist tradition, it has been identified with a method of verification, again by philosopher after philosopher. Nor have these misconceptions had the virtue of exclusiveness; not a few philosophers have held that meaning = method of verification = necessary and sufficient condition.

On the other side, it is amazing how weak the grip of the facts has been. After all, what have been pointed out in this essay are little more than home truths about the way we use words and how much (or rather, how little) we actually know when we use them. My own reflection on these matters began after I published a paper in which I confidently maintained that the meaning of a word was “a battery of semantical rules,”¹⁴ and then began to wonder how the meaning of the common word “gold” could be accounted for in this way. And it is not that philosophers had never considered such examples: Locke, for example, uses this word as an example and is not troubled by the idea that its meaning is a necessary and sufficient condition!

If there is a reason for both learned and lay opinion having gone so far astray with respect to a topic which deals, after all, with matters which are in everyone’s experience, matters concerning which we all have more data than we know what to do with, matters concerning which we have, if we shed preconceptions, pretty clear intuitions, it must be connected to the fact that the grotesquely mistaken views of language which are and always have been current reflect two specific and very central philosophical tendencies: the tendency to treat cognition as a purely *individual* matter and the tendency to ignore the *world*, insofar as it consists of more than the individual’s “observations.” Ignoring the division of linguistic labor is ignoring the social dimension of cognition; ignoring what we have called the *indexicality* of most words is ignoring the contribution of the environment. Traditional philosophy of language, like much traditional philosophy, leaves out other people and the world; a better philosophy and a better science of language must encompass both.

NOTES

1. This is discussed by Ziff, *Understanding Understanding* (Cornell University Press, 1972), especially chapter VIII.
2. This tradition grew up because *the* term whose analysis provoked all the discussion in medieval philosophy was the term "God;" and the term "God" was thought to be defined through the conjunction of the terms "Good," "Powerful," "Omniscient," etc.—the so-called "Perfections." There was a problem, however, because God was supposed to be a Unity, and Unity was thought to exclude His essence being complex in *any* way—i.e., "God" was defined through a conjunction of terms, but God (without quotes) could not be the logical product of properties, nor could He be the unique thing exemplifying the logical product of two or more *distinct* properties, because even this highly abstract kind of "complexity" was held to be incompatible with His perfection of Unity. This is a theological paradox with which Jewish, Arabic, and Christian theologians wrestled for centuries (e.g., the doctrine of the Negation of Privation in Maimonides and Aquinas). It is amusing that theories of contemporary interest, such as conceptualism and nominalism, were first proposed as solutions to the problem of predication in the case of God. It is also amusing that the favorite model of definition in all of this theology—the conjunction-of-properties model—should survive, at least through its consequences, in philosophy of language until the present day.
3. Rather, they will report: "On Twin Earth (*the Twin Earthian name for Terra*—H.P.), the word 'water' means H₂O."
4. The substance of this section was presented at a series of lectures I gave at the University of Washington (Summer Institute in Philosophy) in 1968, and at a lecture at the University of Minnesota.
5. See my 'Is Semantics Possible,' *Metaphilosophy*, 1, no. 3 (July 1970).
6. This assumption is not actually needed in what follows. What *is* needed is that the same *natural kind* can exist in more than one possible world.
7. These points were made in my 1968 lectures at the University of Washington and the University of Minnesota.
8. See Kripke's 'Identity and Necessity', in M. Munitz, ed. *Identity and Individuation* (New York University Press, 1972), p. 157.
9. See my "It Ain't Necessarily So," *Journal of Philosophy* 59 (1962):658–71.
10. See Katz, "Logic and Language: An Examination of Recent Criticisms of Intentionalism," in K. Gunderson, ed., *Language, Mind, and Knowledge* (University of Minnesota Press, 1975).
11. The idea of a "one-criterion" word, and a theory of analyticity based on this notion, appears in my "The Analytic and The Synthetic," in H. Feigl and G. Maxwell, eds., *Minnesota Studies in the Philosophy of Science*, vol. 3 (University of Minnesota Press, 1962).
12. This example comes from an analysis by Anthony Kroch (in his M.I.T. doctoral dissertation, 1974, Department of Linguistics).
13. I *don't* have in mind the Flewish notion of "paradigm" in which any paradigm of a *K* is *necessarily* a *K* (in reality).
14. 'How Not to Talk about Meaning', in R. Cohen and M. Wortofsky, eds., *Boston Studies in the Philosophy of Science*, vol. 2 (Humanities Press, 1965).