

PUTNAM, H.
IN: JOURNAL OF PALEONTOLOGY 70 (1973)
REFERENCE 699

MEANING AND REFERENCE *

UNCLEAR as it is, the traditional doctrine that the notion "meaning" possesses the extension/intension ambiguity has certain typical consequences. The doctrine that the meaning of a term is a concept carried the implication that meanings are mental entities. Frege, however, rebelled against this "psychologism." Feeling that meanings are *public* property—that the *same* meaning can be "grasped" by more than one person and

*To be presented in an APA symposium on Reference, December 28, 1973. Commentators will be Charles Chastain and Keith S. Donnellan; for Donnellan's paper, see this JOURNAL, this issue, 711-712; Professor Chastain's comments are not available at this time.

A very much expanded version of this paper will appear in volume 7 or 8 of Minnesota Studies in the Philosophy of Science (edited by Keith Gunderson), under the title "The Meaning of Meaning."

by persons at different times—he 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).

Secondly, 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. 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* whether a given thing falls into the extension or not). So theory of meaning came to rest on two unchallenged assumptions:

(1) 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 belief are “psychological states”); no one thought that knowing the meaning of a word was a continuous state of consciousness, of course).

(2) That the meaning of a term 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.

ARE MEANINGS IN THE HEAD?

For the purpose of the following science-fiction examples, we shall suppose that somewhere 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, those who call themselves “Americans” and those who call themselves “Canadians” and those who call themselves “Englishmen,” etc.) speak English, there are, not surprisingly, a few tiny differences between the dialects of English spoken on Twin Earth and standard English.

One of the peculiarities of Twin Earth is that the liquid called “water” is not H_2O but a different 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. 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 space ship 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 space ship will report somewhat as follows.

“On Twin Earth the word ‘water’ means XYZ.”

Symmetrically, if a space ship 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 H_2O , and the Twin Earthian space ship will report:

“On Earth the word ‘water’ means H_2O .”

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 $water_{Tm}$ 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” simple isn’t water. The extension of ‘water’ in the sense of $water_{E}$ is the set of all wholes consisting of H_2O molecules, or something like that; the extension of water in the sense of $water_{Tm}$ is the set of all wholes consisting of XYZ molecules, or something like that.

Now let us roll the time back to about 1750. The typical Earthian speaker of English did not know that water consisted of hydrogen and oxygen, and the typical Twin-Earthian speaker of English did

not know that "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.¹

But, it might be objected, why should we accept it that the term 'water' had the same extension in 1750 and in 1950 (on both Earths)? Suppose I point to a glass of water and say "this liquid is called water." 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 same₁ 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—unknown to me—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 sufficient condition: the necessary and sufficient condition for being water is bearing the relation *same₂* 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 *same₂* 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. Thus, the fact that an English speaker in 1750 might have called XYZ "water," whereas 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

¹ See fn 2, p. 710 below, and the corresponding text.

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 *same₁* to the liquid in Lake Michigan, whereas in 1800 or 1850 we would have known that it did not.

Let us now modify our science-fiction story. I shall suppose that molybdenum pots and pans *can't* be distinguished from aluminum pots and pans save by an expert. (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*. If a space ship 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.

This example makes the same point as the preceding example. 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 states 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). 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, suppose I have a *Doppelgänger* on Twin Earth who is molecule for molecule "identical" with me. 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 *heads*!

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 every one 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 every

one 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 every one 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: every one to whom gold is important for any reason has to *acquire* the word 'gold'; but he does not have to acquire the *method of recognizing* whether something is or is not gold. He can rely on a special subclass of speakers. The features that are generally thought to be present in connection with a general name—necessary and sufficient conditions for membership in the extension, ways of recognizing whether something is in the extension, 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 whether 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 before the rise of chemistry. Today it is obviously necessary for every speaker to be able to recognize water (reliably under normal conditions), and probably most adult speakers even know the necessary and sufficient condition "water is H₂O," but only a few adult speakers could distinguish water from liquids that superficially resembled water. In case of doubt, other speakers would rely on the judgment 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, although unknown to almost all speakers who acquire the word. It seems to me that this phenomenon of division of linguistic

labor is one that 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, it 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 is easy to see how this phenomenon accounts for some of the examples given above of the failure of the assumptions (1 and 2). When 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 RIGIDITY

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 *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.

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." 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). On this theory, 'water' means the same in W_1 and W_2 ; it's just that water is H_2O in W_1 , and water is XYZ in W_2 .

(2) One might hold that water is H_2O in all worlds (the stuff called "water" in W_2 isn't water), but 'water' doesn't have the same meaning in W_1 and W_2 .

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. On 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 this 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 a cross-world relation: just understand it so that, e.g., if x is an individual in a world W_1 who is 5 feet tall (in W_1) and y is an individual in W_2 who is 5 feet tall (in W_2), then the ordered pair $\langle x, y \rangle$ 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 $\langle x, y \rangle$ that constitutes an element of the extension of *same height* as, but rather the ordered pair $\langle x\text{-in-world-}W_1, y\text{-in-world-}W_2 \rangle$.)

Similarly, we can understand the relation *same_E* (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_E* 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_E* (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 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_E* 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_E* 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_E* 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* .

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 possible! Conceivability is no proof of possibility.

Kripke refers to statements that are rationally unreviseable (assuming there are such) as *epistemically necessary*. Statements that 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.

In this paper, our interest is in theory of meaning, however, and not in theory of necessary truth. Words like 'now', 'this', 'here' have long been recognized to be *indexical*, or *token-reflexive*—i.e., to have an extension which varies 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, 'T', has two different extensions in two different idioms; 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 *same₂* 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 (2) intension determines extension—cannot be true of natural-kind words like 'water' for the same reason it cannot be true of obviously indexical words like 'T'.

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 'T' in different idioms—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.²

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 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

² Our reasons for rejecting the first option—to say that 'water' has the same meaning on Earth and on Twin Earth, while giving up the doctrine that meaning determines reference—are presented in "The Meaning of 'Meaning.'" They may be illustrated thus: Suppose 'water' has the same meaning on Earth and on Twin Earth. Now, let the word 'water' become phonemically different on Twin Earth—say, it becomes 'quaxel'. Presumably, this is not a change in meaning *per se*, on any view. So 'water' and 'quaxel' have the same meaning (although they refer to different liquids). But this is highly counterintuitive. Why not say, then, that 'elm' in my idiom has the same meaning as 'beech' in your idiom, although they refer to different trees?

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, and this actual nature is not, in general, fully known to the speaker. Traditional semantic theory leaves out two contributions to the determination of reference—the contribution of society and the contribution of the real world; a better semantic theory must encompass both.

Harvard University

HILARY PUTNAM