KIND WORDS AND UNDERSTANDING

MARK PLATTS
Birkbeck College, London, y
Instituto de Investigaciones Filosóficas
UNAM

1. The importance of natural kind words for our understanding of semantics has been urged by Hilary Putnam upon a number of occasions.[1] Those urgings are characteristically imaginative and suggestive; but they take place within what seems to me a quite inadequate theoretical framework. General presuppositions as to the nature or point of meaning-specifications are left unarticulated and so undefended. Too many crucial questions about the language of natural kinds are passed over; indeed, given the almost wilful vagueness of Putnam's approach, many such questions can scarcely even be raised within the terms of his discussion. That same vagueness masks deeply mistaken views about the nature of an ordinary speaker's understanding of his natural language.

In this paper I try to substantiate these sweeping claims through a consideration of one specific thesis about the semantics of natural kind words which Putnam has repeatedly defended and emphasized. If the reasons I give for rejecting that thesis are correct, then the inadequacy of Putnam's framework will have been shown. If, apart from the reasons I give, I am anyway right in rejecting that thesis, important consequences follow for many other aspects of Putnam's treatment of natural kind terms. But this last point will not be explored here.

2. I begin with a brief *reconstruction* of the main (pertinent)

[1] Most notably in the papers reprinted as chapters 8, 11, and 12 of his *Mind, Language and Reality: Philosophical Papers*, Vol. 2 (C.U.P., 1975), to which subsequent page references are made.

points in Putnam's discussion of natural kind words. The ways in which I modify or extend Putnam's own descriptions do not, I think, give rise to any of the objections to that account which I shall make later.

Things are harmlessly simplified if attention is restricted to "clear" cases of natural kind words: 'water', 'gold', 'tiger', 'lemon'. A persistent tradition has sought for, and offered, *a priori* definitions of such words. These offerings are definitions in that they purport to give both necessary and sufficient conditions for the application of the words concerned —true application, not merely justified application. The offerings are *a priori* in the (slightly special) sense that the putative truths stated by the definitions are supposed to be known by anyone who understands the natural kind words concerned in virtue of that understanding: knowledge of the putative truths stated *constitutes* (at least in part) that understanding. Fair examples of such traditional *a priori* definitions are these:

> *water*: colourless, thirst-quenching, liquid, etc.
> *gold*: yellow, heavy, metal, etc.
> *tiger*: striped, fierce, man-eating, animal, etc.
> *lemon*: yellow, tart taste, fruit, etc.

There are at least four kinds of problems facing this tradition which are clearly exemplified in these specimen offerings of that tradition. But the problems transcend the particular specimen definitions just given; and they interact in such a way that they cannot all simultaneously be overcome by tampering with this kind of approach. The interactions suggest that, for natural kind words, there is an incompatibility between the kind of definition required by the traditional approach, one giving necessary and sufficient true application-conditions, and the constraint that such definitions express *a priori* truths (in the sense given).

(i) The first problem is that each definition includes what I shall call, intuitively and without explanation, at least one

higher-level natural kind word: 'liquid', 'metal', 'animal', and 'fruit' in the examples given. Dropping such terms from the examples would emphasize the next two difficulties; yet it is unclear how such higher-level kind words are themselves to be defined within this traditional framework.

(ii) Each specimen definition includes an *etc.*; as such, they fall well short of providing in any interesting sense truly sufficient conditions. But just dropping the *etc.* aggravates the third difficulty, while trying to put something better in its place heightens the fourth.

(iii) It is clear that if we pass over the *etc.*, then one of the definitions proffered provides sufficient conditions for the true application of the term concerned. The point of concocting fools' gold was that it did indeed satisfy the conditions explicitly given; and it would continue to do so however much we may add to the definitions given as long as these additions have any plausible claim to meet the requirement that the resulting definitions express *a priori* truths (in our special sense). But fools' gold is not gold, and there could be fools' water, fools' tigers, and fools' lemons: not just anything which looks like gold, water, a tiger or a lemon has to be gold, water, a tiger or a lemon, even when we admit the higher-level natural kind words mentioned in the definitions to be part of the appropriate look.

(iv) Finally, the definitions given do not provide necessary true application-conditions either. Here the difficulties differ from case to case. Frozen or gaseous water is still water; pure gold is not yellow; an albino tiger is yet a tiger; a lemon is a lemon even when unripe. The proffered definition of 'water' merely describes (part of) the *normal* state and properties of pure water; that of 'gold' describes (part of) how gold *normally* looks (amongst other things); that of 'tiger' describes (in part) the *typical* tiger; and that of 'lemon' describes (amongst other things) part of how lemons *normally*

look in some parts of the world (the normal lemon in Mexico is green).

3. The system in those failings seems obvious, even if we have not articulated its grounds. Putnam describes an alternative account —a radically different kind of account— of how natural kind words are to be "defined". I shall describe that account with considerable oversimplification, while trying to remain closer than in the foregoing to Putnam's own terms of description.

First, there is an account of how natural kind words are *introduced* into a language. Introduction of such a word begins with a reference to some set of paradigm-examplars of the kind concerned: we demonstrate, for example, some typical tigers, some normal-looking bits of gold. To say that the sample is constituted of paradigm-examplars is not to say that of necessity each item in the set distinguished is indeed a member of the kind concerned. There are good questions about what happens in cases of false presupposition of sameness of kind, with a great diversity of such cases to be considered. But let us for present purposes take the simplest case where things do not go wrong at all, where the members of the paradigm-set do all in fact *turn out* to be members of one kind.

In the context of this introduction of the kind word we then hold, for example, that anything is a tiger which is the *same kind of animal* as those in our paradigm-set, or that anything is gold which is the *same kind of metal* as that exemplified by the members of our paradigm-set. If the context makes of the animality or of the metality of the exemplars an obvious feature, then we can simply say that anything is a tiger which is of the *same kind* as the items in our paradigm-set or that any object is gold if it is of the *same* kind as those objects in our paradigm-set.

This reference to *sameness of kind* can then be eliminated —and errors in the supposition of sameness of kind detected— by scientific investigation of the (presumed common) *explanatory structures* within the paradigms: presumably,

atomic structure (at least) in the case of gold and genetic structure (at least) in the case of tigers. That structure, along with other information, will serve maximally to explain the common "surface" properties of the examplars: the stripedness of tigers, the yellow heaviness of gold, for example. (Other information may be used in giving such explanations —say, the presence throughout our samples of gold of copper impurities.) We can then say, for example, that something is gold if and only if it has the same atomic structure as that found in our paradigm-exemplars of gold. And we can then substitute the details of the best available scientific theory as to what that atomic structure is. Those details, obtained through empirical scientific investigation, are obviously only known *a posteriori* in the standard sense; so the statement of the membership-conditions of the natural kind which issues from the substitution of those results will also be a (putative) *a posteriori* truth in the standard sense.

4. I assume that skeleton account just given to be correct, albeit wildly oversimplified, for at least some favourable cases of relatively "untheoretical" natural kind words. Some further comments will both eliminate pertinent oversimplifications and prepare the ground for later discussion.

Let me say, at this point without explanation, that the *austere sense* of a natural kind word is given by a statement with a form something like, say,

(AS) 'is gold' is true of α iff α is gold

(The *something like* is returned to later.) Why have such *austere rules* not been accepted as plausible candidates for the traditional *a priori* definitions (in our special sense of *a priori*)?[2]

[2] This (apparent) possibility is what explains my saying that the problems (i)-(iv) outlined in § 2 *suggest* an incompatibility between the kind of definition required by the traditional approach and the constraint that such definitions express *a priori* truths. But I am not now, or later, defending this possibility for it rests upon a supposition of (*a priori*) knowledge of rules specifying *word* meanings which I would reject.

Many quite distinct reasons can be historically discerned (a history that lives on) as being (generally tacitly) operative. One thought is that such rules are trivially true and so express trivial truths. That seems to have been Putnam's thought; I return to it later. More pressing at this point are other reasons for rejecting such austere rules as possible descriptions of the meanings of natural kind words, reasons which reflect (tacit) constraints, other than that of non-triviality, upon acceptable rules for describing meanings.

One such thought is that these austere rules could not be used to impart mastery of the words quoted (designated) to one who previously lacked such mastery. Another, more metaphorical, thought is that such austere rules give no decompositional analysis of the sense of a natural kind word, give no indication of the composition of that sense. The non-metaphorical content of this idea is that an austere rule gives no guide as to which are the sound inferences essentially involving a given natural kind word. A yet further thought is that an austere rule could not function as a guiding-rule, as one which guides usage of the word concerned. And, finally, there is the thought that an austere rule gives no guidance as to how statements using the kind word concerned are to be verified, in however weak but significant a sense of verification.

That is a very mixed bag of considerations. Let me risk confusion by lumping those considerations together in the following way. By a statement of the *definitional sense* of an expression I shall understand a statement of its necessary and sufficient true application-conditions with as least one of the following properties:

(i) it is at least possible that the statement serve, for some speaker, to impart to that speaker mastery of the expression concerned; *or*
(ii) the statement is "decompositional" in the sense that it is at least possible that it account for some of the inferences involving statements in which the word concerned occurs; *or*

(iii) it is at least possible that the statement serve, for some speaker, as a guiding-rule in the usage of the word concerned; *or*
(iv) it is at least possible that the statement serve as a guide as to how statements containing the word concerned should be verified.

One reason for constructing such a confusion-inviting definition is this: if Putnam's account of natural kind words is accepted, this has an important bearing within a wide range of ideas about philosophical meaning-analysis. According to that account, a statement of the definitional sense of a natural kind word will be obtained only through scientific investigation of the members of the paradigm-set (remember that necessary and sufficient true application-conditions are required). That is: the reference (extension) of a natural kind word is (partially) prior to its definitional sense; only after we have (partially) fixed the reference of the word concerned can we go on to discover its definitional sense.

That is important because it would show the inapplicability in the case of natural kind words of an element to be found, with differing degrees of explicitness, within many different accounts of meaning-analysis. This is the idea that the definitional sense of an expression —where that is understood according to at least one of the disjuncts (i)-(iv)— can be known prior to having any knowledge of its extension. A limiting case of this idea is the supposition that the definitional sense of an expression can be known even when that expression has no extension (is "empty"). Natural kind words, on Putnam's account, show that as a general thesis without restriction such an idea is indefensible.

There is another reason for presenting this monstrosity of a definition of the definitional sense of an expression. Putnam's own views as to the adequacy constraints upon philosophical analyses of meaning, upon meaning-specifications, are far from clear. At different points in his discussion he is apparently influenced by various of the possible constraints

lumped together in the definition given. The pejoratively *ad hoc* character of that definition is meant to mirror the similarly *ad hoc* character of Putnam's (tacitly operative) ideas about meaning analysis.

5. Another preliminary set of remarks. I said that any statement of the definitional sense of a natural kind word will be a (putative) *a posteriori* truth, the result of empirical investigation of the partially fixed extension. But might not these (putative) truths yet be *a priori* in our special sense? Might not the definitional sense of a natural kind word be known by anyone who understands that word in virtue of that understanding?

The immediate intuitive answer is negative. The definitional sense of a natural kind word, the requisite set of necessary and sufficient true application-conditions, is given using the best available scientific theory of the kind concerned. There seems no reason to suppose that the ordinary speaker, while competent in the usage of the word concerned, need have the least idea of what that theory is.

If that is so, then Putnam's account has as a consequence that an ordinary speaker, while competent in the usage of some natural kind word, yet need not know how to define the word in the sense of being able to state its definitional sense. But why should that worry? If we were to assume (a) that the definitional sense it at least "part of" the meaning of a natural kind word; (b) that anyone competent in the usage of a word knows its meaning; (c) that the knowledge so possessed can be stated by the speaker; and (d) that ordinary, non-expert people can be competent in the usage of natural kind words; then, of course, the account given by Putnam would be wrong. But it is difficult, to say the least, to specify ways of construing 'the meaning of a word' and 'knowing the meaning of a word' which leave each of (a), (b), and (c) plausible, and which leave the argument leading to the rejection of Putnam's account free of fallacious word play.

Putnam makes an addition to the account just sketched which can be construed in the present context as a way of making the denial of either (b) or (c) palatable once the argument is freed of punning. (His attention seems to be focused upon (b); but one with no truck for what Putnam calls 'methodological solipsism' (*op. cit.*, pp. 219f) might shift attention to (c).)

How can someone be competent in the usage of a natural kind word whilst having no idea of what the definitional sense of that word is —the definitional sense which specifies the true application-conditions of that word? Won't his employment of the word be arbitrary, wholly without justification? Such a speaker will have no way of settling any disputed cases which crop up. But worse, he cannot justify his usage even in undisputed cases: nothing he can say serves to show that such cases are beyond dispute. An answer to this line of thought which can be read into Putnam's work begins from the idea that the ordinary, non-expert speaker is a member of a linguistic community which includes some people, the experts, who can indeed state the best available scientific theory of the particular kind concerned. In virtue of this, the ordinary speaker does have a way, for example, of settling any disputed cases: he can resort to the experts. And it is this same access to expert opinion which frees his ordinary usage of the kind word concerned in "unproblematic" cases from the arbitrary: the justification for that everyday usage is that were any (previously unproblematic) case to be disputed, there exists a means available to him of deciding the case.[3]

Putnam is thus led to propose his 'hypothesis of the universality of the division of linguistic labour' whereby 'every linguistic community . . . possesses at least some terms whose associated "criteria" [roughly, our definitional sense] are known only to a subset of the speakers who acquire the terms,

[3] On the question of whether the ordinary speaker's usage would be *explicable* (not justifiable), the discussion of the unreflective nature of language use in § 14 has a direct bearing.

and whose use by the other speakers depends upon a structural cooperation between them and the speakers in the relevant subsets' (*op. cit.*, p. 228).

In a community where such a structure exists, it need not matter that the ordinary speaker cannot state the definitional sense of a natural kind word he uses; indeed, it would be an enormous hindrance were he required to do so before being "allowed" to *use* the word. In such a community the rejection of (b), in any sense in which it can count against Putnam's theory, should be quite unworrying.

It is important to distinguish clearly between this "sociolinguistic" phenomenon of the division of linguistic labour and the "logico-semantic" feature of natural kind words discussed earlier —the partial priority of reference *vis à vis* definitional sense, what Putnam calls the 'indexicality' of natural kind words. The two features are quite independent; the existence of legal experts testifies to this in one direction, the possibility of a community consisting just of expert physicists in the other.

6. What is the meaning of a natural kind word? What is the general 'normal form' for specifying the meaning of any such word? Putnam develops a subtly complicated answer to these questions; his final resting-point is exemplified by the following 'normal form description' of the meaning of 'water' (*op. cit.*, p. 269):

Syntactic Markers	*Semantic Markers*	*Stereotype*	*Extension*
mass noun; concrete	natural kind; liquid	colourless; transparent; tasteless; thirst-quenching; *etc.*	H_2O (give or take impurities)

That point is reached at the end of a long paper in which Putnam declares his interest to be 'in theory of meaning'. But

there is a striking *lacuna* in that paper: any consideration of the general nature of a theory *of* meaning, of the theoretical *point* of the notion of meaning itself. Perhaps this omission is intentional. Putnam tells us that the 'amazing thing about the theory of meaning is how long the subject has been in the grip of philosophical misconceptions ... and how weak the grip of the facts has been' (*op. cit.*, p. 271). But questions remain as to which "facts" bear upon the notion of meaning and as to quite how they do so. Answers to those questions will require thought as to the point of the notion of meaning, its theoretical role within an overall theory of language use.

7. The unsatisfactoriness of Putnam's way with meaning emerges clearly upon consideration of his reasons for including as "part of the meaning" of a natural kind word what he calls its *stereotype*. This is, roughly, the description of more or less obvious (more or less "observational") properties of the typical or normal members of the kind concerned (cp. the last paragraph of § 2). There are difficulties about the definition of the notion of a stereotype —for example, that of explaining the difference between stereotypical properties and those mentioned under 'Semantic Markers'— but they are not the present concern.[4]

According to Putnam (but in my terminology), what traditional accounts have offered as the definition of a natural kind word is just the stereotypical description along with some of the "semantic markers". That seems (roughly) right. But Putnam's retention of the stereotype as still part of the meaning-specification stands in need of reasoned support if the suspicion is not to linger that its presence there is but the residual ghost of the *a priori* model of meaning-analysis which Putnam himself has so compellingly criticized.

8. In 'Is Semantics Possible?', Putnam argues that 'there

[4] Cp. *op. cit.*, p. 230; pp. 249-52; pp. 266-8.

are a few facts about "lemon" or "tiger" . . . such that one can convey the use of "lemon" or "tiger" by simply conveying those facts. More precisely, one cannot convey the approximate use *unless* one gets the core facts across' (*op. cit.*, p. 148). These core facts include the stereotypical description.[5] And in 'The Meaning of "Meaning"' Putnam uses this same idea when attempting to explain the notion of a stereotype through that of what it is *linguistically obligatory* to convey when instructing someone in the usage of a natural kind word (*op. cit.*, pp. 249-52).

One obscurity here is obvious enough: the connections presumed to obtain between "conveying the approximate use" on the one hand, and a theoretical representation of the meaning on the other, are left unarticulated and undefended. (The idea of *giving the meaning* of a word is full of traps for the unwary.) Putnam, it seems, is subscribing to one of the traditional ideas about meaning-analysis mentioned earlier, namely, that any rule purporting to describe the meaning of an expression must be one that could be used to impart mastery of the expression to one who previously lacked it (meaning as a *gift*; cp. § 4). That is a clearish constraint, but Putnam does not defend it. As a completely general requirement upon 'normal form description' of meaning it is incoherent; Putnam gives no reason to retain it in the special case of natural kind words.

Any thesis about what *has* to be done in "conveying the use" of a word looks to be a piece of transcendental (armchair) genetic psychologising. It appears anyway that Putnam's thesis cannot be that, in teaching the usage of a natural kind word, one has to *state* the stereotypical description; for he seems to hold that one 'way of telling someone what one means by a natural-kind term' is 'so-called ostensive definition', as in 'this (liquid) is water', 'this (animal) is a tiger', and so on.[6] (Remember that *liquid* and *animal* fall under semantic markers, not the stereotypical description.) So the

[5] *Op. cit.*, p. 148 and pp. 150f.
[6] *Op. cit.*, p. 229.

thesis must be that one has to 'get across' the stereotypical description, whether one states it or not: the audience has somehow to divine the stereotype from the samples demonstrated in cases where that stereotype is not stated.

The gap in these transcendental ruminations is obvious. Let us concede, for present purposes, that if the audience is to acquire even approximate understanding of the word concerned through our ostensive definition, then they must form some pertinent, plausible beliefs about the kind (partially) demonstrated. Let us even concede that the 'pertinent' here means that the beliefs should be about (presumed) typical features of the kind concerned. Still, it does not follow that there is some one set of beliefs about some one set of typical features which must be got across. Typical tigers have ever so many features in common. Some, for varying reasons, may usually be more striking to us. But that does not imply that one who has, say, eccentric interests cannot come to acquire the use of the word 'tiger' through ostensive definition just as well as we do; all that is needed is that his eccentric interests still direct him to typical features of tigers.

It is a different question as to whether the beliefs concerned need to be about typical features at all. Isn't it enough that the beliefs our learner so acquires be intelligible? One (too) quick a way of rejecting such a possibility is the following: to take over another of the traditional ideas about meaning-analysis, namely that meaning-specifications must be such that they could guide the speaker in his usage of the word concerned (cp. § 4), along with the idea that such guidance must at least be sufficient for the typical ("paradigmatic") cases of the kind concerned. This is just too quick a way of eliminating apparent possibilities. Suppose a group of people live in conditions which they have general reason to believe are likely to produce untypical animals, and so untypical tigers; but their grounds for that belief do not equip them to "construct" the details of the typical tiger. Couldn't they still acquire, and transmit to the young, a perfectly intelligible use of the word 'tiger'? They might have

a different *conception* of what tigers are; that does not yet show them to have a different *concept*. I return to this point later (§ 12).

9. In 'The Meaning of "Meaning"' another instructive argument for the retention of the stereotype is presented. Putnam says that '... nouns like "tiger" or "water" are very different from proper names. One can use the proper name "Sanders" correctly without knowing anything about the referent except that he is called "Sanders" —and even that may not be correct ... But one cannot use the word "tiger" correctly, save *per accidens*, without knowing a good deal about tigers, or at least about a certain conception of tigers. In this sense concepts *do* have a lot to do with meaning' (*op. cit.*, p. 247).[7] The supporting argument comes a page later: 'Suppose our hypothetical speaker points to a snowball and asks, "is that a tiger?". Clearly there isn't much point in talking tigers with *him*. Significant communication requires that people know something of what they are talking about ... What I contend is that speakers are *required* to know something about (stereotypical) tigers in order to count as having acquired the word "tiger"... The nature of the required minimum level of competence depends heavily upon both the culture and the topic ...' (*op. cit.*, pp. 248-9).

It is quite unclear what qualification to his thesis Putnam wishes to introduce with his final remark, so I shall pass it over. At least one of our previous doubts immediately applies to this new argument. Even if 'speakers are *required* to know something about (stereotypical) tigers in order to count ..., etc.' it does not follow that there is something, some one thing or some one set of things, that they are '*required*' to know. But deeper worries arise about the ground of the requirement, about whether what is required is knowledge,

[7] On what has to be known by one who competently uses a proper name, compare Putnam's remarks at p. 201 and p. 203 where knowledge that 'Quine' refers to a person is said to be necessary. Note also that the *per accidens* in the passage cited in the text invites the thought that a conception of rules of language as guiding-rules is influencing Putnam (cp. § 4).

and, again, about whether the objects of this "knowledge" must be stereotypical tigers.

The aim of a theory of interpretation is to make sense of a person's linguistic behaviour as part of making sense of him. The constraint upon any such theory is that it make sense to say of the person that he was then and there saying what our theory represents him as saying. The applicability of that constraint is determined by a further one: namely, that the propositional attitudes which we are led (in diverse and complex ways) to attribute to the speaker as a result of our representation of what he was saying can intelligibly be attributed to him. That intelligibility is itself controlled by all we know and believe about the speaker, including things that we know or believe as the result of our interpretation of his other linguistic actions. (I here telescope a long, and I hope familiar, line of thought; this telescoping continues into the next paragraph.)

Central to the attitudes so ascribed, and so constraining our theory of interpretation, are beliefs. So, for example, if we are to treat someone's emissions of 'tiger' as manifesting understanding of that word, as competent uses of that word, we must be able intelligibly to attribute to him some beliefs about tigers. But what have to be attributed are beliefs, not knowledge: false but intelligibly acquired beliefs would suffice, given the aim of our theory of meaning. There may be graver difficulties about attribution of false beliefs about stereotypical matters, but examples are easily imaginable in which such beliefs are precisely those which should be attributed. Nor is it clear that such beliefs must be about stereotypical matters according to us, according to what strikes us. Indeed, it is not even clear that the beliefs attributed must include some which our subject believes to be about stereotypical matters; maybe he knows (or believes) that he lives in conditions which are likely to produce freaks.

Most important of all is that the ground of our requirement is utterly different from that which Putnam envisages. It has nothing to do with what is required or obligatory in

the teaching of a natural kind word. It has nothing to do with the requirement that 'people know something of what they are talking about'. It has everything to do with quite general considerations about the role of a theory of interpretation, a role which fixes the theoretical point of the notion of meaning itself.

In this context the oddity of Putnam's remarks about proper names can be seen. 'One can use the proper name "Sanders" without knowing anything about the referent except that he is called "Sanders" and even that may not be correct.' But suppose our hypothetical speaker points to a snowball and asks 'is that Sanders?'. Clearly there isn't much point in talking Sanders with *him*. 'Significant communication requires that people know something of what they are talking about.'

Here Putnam sees that knowledge is not at issue —the one belief about the referent can be false. Yet his position is wide open to an argument *ad hominem* of the form *ad absurdum*. How did he come by such a position? Because, I surmise, he has bought Kripke's well-known arguments against "description theories" of the meaning of proper names (cp. *op. cit.*, pp. 203-4), and yet has failed to see the crucial role of belief-ascription within a structured theory of interpretation. Admitting the role of belief-ascription within the general theory of language use does not *yet* commit us to any kind of "description account" of the meaning of proper names *within* the (component) theory of meaning itself. But the same moral carries over to "description theories" of natural kind words, even when weakened from the traditional model to Putnam's far more sophisticated (albeit hybrid) account.

10. In 'The Meaning of "Meaning"' there is the hint of another argument which may have influenced Putnam; it could easily influence others. It begins with the question as to how the meaning of a natural kind word should be described. One possibility, introduced earlier, is use of an austere rule like

(A_1) 'is water' is true of α iff α is water

I charted some of the constraints one might impose upon meaning-specifying rules which would rule out such austere candidates (§ 4). Putnam's own objection to austere rules is clearly stated: they 'don't say anything about the meaning of the word "water"' (*op. cit.*, p. 261). Suppose, then, we try instead a rule like.

(A_2) 'is water' is true of α iff α is H_2O

Putnam's view of such rules is that 'what they say is wrong as a description of the *meaning* of the word "water"' (*op. cit.*, p. 261). And so it seems, in Putnam's words, we are 'left with nothing'.

But what exactly is Putnam's ground for rejecting (A_2)? Taking a hint from a later remark (*op. cit.*, p. 269), the idea might be that acceptance of (A_2) as the specification of the meaning of 'is water' would 'mean that knowledge of the fact that water is H_2O is being imputed to the individual speaker or... to the society'. While (A_1) is uninformative about the *meaning* of the *word* 'water', (A_2) is too informative about what water *is*. Usage of (A_2) as a specification of meaning would falsify originary speakers' *understanding* of their language by falsifying their knowledge (i.e. ignorance). What is need is an "informative" rule which does connect with ordinary speakers' understanding, with what ordinary speakers *know*. How else could this need be satisfied except by the (supposedly universally known) stereotypical specification?

However, it is simply false to say that (A_1) says *nothing* about the meaning of 'water'. Such an austere rule, in virtue of the shift from *designation* of an expression on one side of the biconditional to *use* of that same expression on the other side records, in the English language, a (putative) contingent, learnable, forgettable fact about (in this case) the En-

glish language.[8] Which is just the kind of fact a theory of meaning for English should be trying to record in the first place! It could be trivially true (although in fact it is not) that (A_1) expresses a truth; it still would not follow that the truth so expressed is itself trivial.

I said that there is a hint of this argument in Putnam; I am not at all sure that, for him, it is a further argument. Consider the claim that 'what [rules like (A_2)] say is wrong as a description of the *meaning* of the word "water" '. I elaborated upon this in terms of the connections between meaning, understanding, and speakers' knowledge. For that elaboration to work —both as a reason for rejecting (A_2) *and* as a motivation for the search to be ended by discovery of the stereotypical specification— it has to be assumed, not just that any knowledge attribution to speakers consequent upon some determinate meaning-specification must be a defensible attribution, but also that any acceptable meaning-specification will, in virtue of its acceptability, require the attribution of some ("informative") knowledge to those who understand the expression concerned.[9] But that is precisely part of the question at issue: given that competent speakers generally cannot state the rule giving the definitional sense of a natural kind word, is there anything else they must know in virtue of that competence? Must understanding of a word imply some determinate ("informative") cognitive state? That is the question begged in the reconstructed line of thought.

Within my, theory-laden, framework there is no difficulty

[8] Cp. Putnam, *op. cit.*, p. 246 and pp. 270-1. Maybe what Putnam means is that austere rules do not say *enough* about the meaning of natural kind words like 'water'; and he might try to defend such a view by pointing out the inability of such rules to satisfy the constraints upon meaning-specifications incorporated into the definition of definitional sense given in § 4. But then he must say which of those possible constraints he wishes to invoke and why those constraints should indeed be adopted.

[9] There are problems enough with this assumption when the "expression" is a sentence, suitable for performing a complete linguistic action; when, as in the present case, the expression is a word the thesis that understanding its usage requires propositional knowledge (even austerely specified propositional knowledge) of the rule describing that usage is indefensible. Cp. fn. 2 above and § 14 below.

about explaining the unacceptability of (A_2). The acceptability of a theory of meaning is determined, in this framework, by its role within an overall interpretative theory of language use which leads to plausible propositional attitude ascriptions. The meaning-specifications pull their weight in specifying the *contents* of propositional attitudes assigned to speakers *via* specification of the contents of their linguistic actions. Use of (A_2) as such a meaning-specification could lead, for example, to the attribution to someone who says 'the water is cold' of the belief that the H_2O is cold. But the attribution to such a speaker of *any* beliefs about molecular structure might be quite without independent foundation, might be (at best) quite gratuitous. And that is why (A_2) must be rejected: not because of some unargued-for assumption of connections between meaning, understanding, and speakers' knowledge, but because of its inadequacy within a theory of meaning given the role of such a theory within an overall account of language use.[10]

11. Even when the illusion of utter triviality about austere rules has been dispelled, an anxiety is wont to remain. Consider the case where we are trying to construct a theory of meaning for English *in* English. Our object language is part of our meta-language, the latter exceeding the former only in certain semantic vocabulary. In such a case, wouldn't employment of austere rules make our task too easy?

Here I shall make three, unduly brief but closely connected, remarks. These remarks are designed to gesture at the true problems facing the construction of a theory of meaning —problems which can be masked by the pursuit of non-austere, "decompositional" meaning rules. Each remark merits a far more detailed elaboration than I can give here.

First, acceptability of a theory of meaning is determined by acceptability of the overall *interpretative* theory of which it is a component. Understanding of what it is for a theory of

[10] Cp. my *Ways of Meaning* (London, 1979), pp. 64-7.

meaning to be a good theory of meaning requires understanding of what it is for a theory of interpretation to be a good theory of interpretation; there is nothing trivial about explaining the latter. This notion of interpretation —and so of a good interpretation— is just as much operative in the English-English case as in, say, the Englis-Spanish case. We who construct the theory of meaning are indeed presumed to understand the language in which we do so (the meta-language); so, in the English-English case, we are also presumed to understand the object-language. But that understanding of the object-language is the perfectly natural, *unreflective* understanding which is attributed to all native speakers of a language which they use competently. It is a practical competence in usage, not a theoretical propositional knowledge of the theory of meaning for the language. In constructing an interpretative theory of meaning we are trying to construct a theoretical (propositional) representation of our customary practical capacities. To be sure, when within the constraints of interpretation in general, we present a theory of meaning including austere rules, there remain questions as to *which* truths are stated by those austere rules; but the idea that there *must* be some non-austere way of saying what those truths are is an idea, reductive in nature, which if adopted as a general constraint is incoherent. For us, there cannot be *in general* a further question.

The other two remarks explain my intermittent use of 'something like' when giving austere rules. Even within the use of a generally austere framework in the English-English case, substantial questions arise as to the "logical grammar" of the expressions in the language under study. For example, many natural kind words are *mass terms*; yet there is to date no plausible account of how such terms are to be handled even within a generally austere framework. We do not know the *logical forms* of all the kinds of sentences in which mass terms can occur; for we do not know the kind of contribution, or kinds of contribution, made to those logical forms by mass term natural kind words. Recent discussion of the

semantics of mass terms shows that there is nothing trivial about constructing a theory of meaning for English even when austerity is in principle accepted. (Comparable worries arise, of course, for natural kind terms which are not mass terms.)

The third point is this. In presenting Putnam's account of natural kind terms I followed him closely in sketching a quasi-historical picture of how such terms are *introduced* into a language. But what matters is to distinguish natural kind terms from other kinds of kind term by reference to their functioning *within* a language. The relevant distiguishing feature —earmarked "historically" by talk of indexicality or of partial priority of reference— is one which has to be explained and detected in disregard of the strictly irrelevant and occasionally misleading matter as to how such terms were first brought into the language.[11] (Of course, functional differences may correlate to some high degree with such "historical" differences.) Suposse we do find some such way of distinguishing natural kind words *in usage*; then a further question immediately arises. Is the distinction so made between natural kind terms and other kind terms one which must receive explicit acknowledgment within our theory of meaning itself? Is the distinction really a *semantic* distinction, one pointing to a difference of kind in contribution to logical forms? It need not be so, for not all differences in usage reflect strict semantic differences; another possibility, for example, is that the presumed difference between natural and other kind words be acknowledged at the point of assessing the plausibility of propositional attitude ascriptions —such plausibility being determined in different ways for natural and other kind words.[12] Decision between these (and other) options would require consideration both of the general aims of interpretation and of the role within interpretative theories of assignments of strictly semantic properties;

[11] Cp. Putnam, *op. cit.*, p. 205.

[12] Cp. Putnam, 'Language and Reality', *op. cit.*, pp. 272-290 for some considerations which could be brought to bear upon this question.

acceptance of a general austerity makes no such decision for us.

12. There is one last argument for the retention of the stereotype to be considered. In 'Is Semantics Possible?', Putnam considers a possibility in which lemons turn blue and "stay that way". He says that 'then in time "lemon" would come to have a meaning with the following representation:

lemon: natural kind word	Associated characteristics: *blue* peel, tart taste, etc.

Then "lemon" would have changed its meaning'.[13] That is: changes in stereotypical description produce changes in meaning, so that the stereotype must be *part of* the meaning.

Two quite different theses must be distinguished. One is that changes in stereotype, in stereotypical beliefs, are relevant to questions about sameness or difference of meaning. The other is that *any* change in the stereotype (however "small") implies a change of meaning (however "small"). The second is what is needed to obtain immediately the result that the stereotype is part of the meaning. But there is all the difference in the world between the two theses.

Let me change the terminology here. (This helps to destroy the illusion of 'home truths'.) Instead of talking about sameness or difference of meaning, I shall talk of sameness or difference of *concepts*. My question is not 'does "lemon" have the same meaning as it did two hundred years ago?', but rather 'does "lemon" now express the same concept as it did two hundred years ago?'. Any threat of Platonism is avoided by construing such questions in terms of interpretation. The question becomes, say: is the following a meaning-axiom in the best interpretative theory (best for us) of English two hundred years ago?

'is a lemon' in English (18th century) is true of α iff α is a lemon?

[13] *Op. cit.*, p. 148; but compare pp. 142-3.

Talk of concepts must be distinguished from talk of *conceptions*. My conception of, say, a lemon is discovered by seeing which properties of objects serve best to explain my identification of given objects *as* lemons (whether they be so or not).[14] In many cases I shall have explicitly held beliefs about what those properties are; when such beliefs are true my stereotype of a lemon will generally constitute at least part of my conception of a lemon. (There are good questions about cases where my beliefs about the grounds of my identification seem at variance with my actual linguistic practice; these too will be passed over here.) The conscious part of a speaker's conception, including his stereotype, will have an important role to play in our understanding of what the *speakers means* by utterances of sentences containing the word concerned.[15]

What, then, are the relations between questions about sameness or difference of concepts and questions about sameness or difference of conceptions? Two limiting theses can be imagined. *Extreme idealism* holds that any difference of conception implies a difference in concepts. *Extreme realism* holds that difference of conceptions has no relevance to questions of sameness or difference of concepts. Both extremities are madness.

Questions of sameness or difference of concepts are settled by questions about what is the correct interpretative theory. Those questions in turn are constrained by the condition of plausibility of propositional attitude ascription, amongst which attitudes are beliefs. Against the extreme idealist thesis, the following seems a clear possibility: that the best in-

[14] Compare Putnam at pp. 196f and at p. 230. What Putnam wants to say in defense of Engel's theories is far better put in this terminology than in Putnam's. Note also the confusion on this point in Putnam's remarks quoted above at p. 19.

[15] Note Putnam's "speaker's meaning" phrasing in the passage quoted at p. 17 above. In what I say in the text I am in conflict terminologically with what Ronald Dworkin says in his *Taking Rights Seriously* (London, 1977) at p. 135 where he *identifies* speaker's meaning and speaker's conception; for me, much of a speaker's conception may be quite unconscious (cp. § 14 below), and quite ill-fitted for any role in an account of what the speakers means.

terpretation of another's usage of the word 'lemon' be obtained by an austere pairing of that expression with my use of the word 'lemon', even though he and I have different conceptions of what lemons are, given a case in which that difference in conception is intelligible to me granted what I know and believe of this situation and mine. But this perspective counts against the extreme realist thesis too. Questions of interpretation (and so of sameness and difference of concepts) are constrained by considerations of plausibility of propositional attitude ascription. An interpretative theory, settling "conceptual" questions, must issue in plausible propositional attitude assignments. If the propositional attitudes assigned under a scheme of sameness of concepts are not intelligibly attributed, then difference of concepts there is. So difference of belief —or *sameness* of belief— stands in intimate connection with matters of sameness or difference of concepts; beliefs ascribed under a scheme of interpretation, whether the same as our own beliefs or not, must be plausibly ascribed if that scheme of interpretation is to be accepted. Since some of a person's beliefs find expression in a part of his conception of the kind concerned, including his stereotypical beliefs, there cannot be the extreme independence of concepts and conception that the extreme realist thesis posits.

There is therefore no tidy formula about the relations between sameness of concept and sameness of conception, for there is no tidy formula for interpretation. Putnam's argument requires a tidy formula —the extreme idealist thesis— at least in the case of natural kind words. But he has given us no reason to accept that tidy formula; nor has he given us untidy considerations which could yield desired result.

13. The general morals to be drawn from the preceding are obvious. Of course the 'grip of the facts' should not be weakened; of course if all we have are *theories* of meaning something has gone badly wrong. But the solution is not to return

to casual, not even casual philosophical, talk about 'meaning' in stating 'home truths' about semantics.[16]

Putnam tells us at one point: 'What we have analysed so far is the predominant sense of natural-kind words (or, rather, the predominant extension). But natural kind words typically possess a number of senses. (Ziff has even suggested that they possess a *continuum* of senses.)' (*op. cit.*, p. 238). Such a claim is as obscure as it is puzzling.

Within the, broadly speaking, Fregean tradition there is much work designed to make of the notion of sense a clear notion. I hold no brief for all of that tradition since much of its shares failings of Putnam's work. But within that tradition one finds, for example, the interesting suggestion that a rule purporting to give the sense of an expression is an adequate specification of that sense only if knowledge of the truth expressed by that rule would suffice for understanding that expression, would suffice for mastery of the use of that expression (granted competence in the use of some other expressions, of course).

Note that this is not the same as the requirement (cp. § 4) that the rule stating the sense be suitable for *imparting* mastery. Austere rules remain plausible candidates.[17] But more importantly, given this relatively exact idea, Putnam has said nothing to show that natural kind words 'typically possess a number of senses'. Nor, given that idea, could any sense be make of talk of the 'predominant sense' being the 'predominant extension' since that merely confuses a notion tied to knowledge of truths with one tied to knowledge of objects. There exists *a* notion of sense in ordinary English, as my last sentence shows; but the bearing of that notion, and of truths stateable using it, upon our theories of meaning, and so upon our theoretical notion of meaning (or sense), is a substantial problem.

[16] But compare Putnam at p. 146.

[17] On sense, austerity and the avoidance of psychologism, see John McDowell, 'On the Sense and Reference of a Proper Name', in Platts (ed.), *Reference, Truth and Reality* (London, 1980), pp. 141-66.

14. One final, tentative moral is less obvious. Putnam rejects the idea that the 'key problem' in semantics is: how do we come to understand a new sentence? (*op. cit.*, p. 149). His view, apparently, is that the question 'how do we come to understand a new *word?*' is far more important.

Frege held that asking for the meaning of a word in isolation, not in the context of a proposition, almost forces us into psychologism, into offence against the principle that we should 'sharply separate the psychological from the logical, the subjective from the objective'.[18] This Fregean thought seems well-exemplified in Putnam's account.[19] His retention of the stereotypical description as part of the meaning of a natural kind word is of a piece with the psychologism in his account of speakers' understanding of such words, that account emphasizing part of the "conscious" part of speakers' conceptions of the kind concerned. This same psychologism is, I surmise, inevitably reinforced by asking, not 'what account can be given of a speaker's understanding of a natural kind word?, but rather 'how does a speaker *come to* understand a natural kind word?', with its presupposition that a meaning-specification must be suitable for imparting mastery to one who previously lacked it.

Psychologism in the theory of understanding falsifies the phenomenology of language use; that falsification receives its "theoretical" mirror-image in psychologistic accounts of sense. The greater our competence in the usage of a term, the less we need to rely upon any view as to its associated "criteria", in however weak a sense, in our application of that term. The greater our competence, for example, the less we need recourse to our (stereotypical) beliefs when making judgements involving that term. At the limit, those beliefs play no role in the phenomenology of our mastery of that expression, in the phenomenology of our understanding of it. Which serves to introduce the possibility, rules out by Put-

[18] *The Foundations of Arithmetic, trans.* J. L. Austin (Blackwell, 1950), at p. x.

[19] Note, especially, Putnam, *op. cit.*, pp. 260-2; but compare also p. 222.

nam's account, of that understanding trascending differences in speakers' (stereotypical) beliefs.

If we remember that much of what passes as "dictionary definition" of a natural kind word is its stereotypical specification, then we find that the point here was well put by Russell in *The Analysis of Mind* (London, 1921): 'Understanding words does not consist in knowing their dictionary definitions, or in being able to specify the objects to which they are appropriate ... Understanding language is more like understanding cricket: it is a matter of habits, acquired in oneself and rightly presumed in others. To say a word has meaning is not to say that those who use the word correctly have ever thought out what the meaning is ...' (pp. 197-8).

I have never fully understood the ground of Frege's conjecture as to how we are 'almost forced' into psychologism by seizing upon the word as the primary unit of meaning. But although there are other sources of Putnam's psychologism, his striking exemplification of Frege's warning might well make one think again upon the ground of that conjecture.

RESUMEN

1. Putnam ha señalado la importancia que tienen las palabras que designan clases para nuestra compresión de la semántica; pero Platts cree que sus formulaciones se presentan en un marco teórico inadecuado, pues muchas preguntas cruciales acerca de tales palabras se dejan de lado. En este artículo el autor propone fundar esta afirmación examinando una tesis específica acerca de la semántica de tales palabras que repetidamente Putnam ha defendido y subrayado.

2. Platts presenta una breve *reconstrucción* de los puntos principales del análisis de Putnam de las palabras para clase, y restringe nuestra atención, por simplicidad, a casos "claros" de palabras para clases naturales, p. ej., "agua", "oro", "tigre", "limón". Una tradición persistente ha buscado y ofrecido definiciones *a priori* de tales palabras. Estas definiciones pretenden dar condiciones necesarias y suficientes para la aplicación de las mismas; y las definiciones son *a priori* en el sentido (ligeramente especial) de que se supone que las verdades putativas ofrecidas por tales definiciones las conoce quienquiera que entienda tales palabras sólo en virtud de que las entiende. Ejemplos de tales definiciones *a priori* serían:

agua: líquido incoloro, quita la sed, etc.
oro: metal amarillo, pesado, etc.

Frente a tal tradición, ejemplificada por estas definiciones, surgen, al menos, cuatro tipos de problemas. Estos trascienden a las anteriores definiciones tipo e interactúan de tal manera que no pueden superarse todos simultáneamente modificando este tipo de enfoque. Los problemas son:

(i) Cada una de las definiciones incluye lo que, intuitivamente, Platts denominará palabras para clases naturales de *nivel superior*: "líquido", "metal", en los ejemplos. ¿Cómo han de definirse estas palabras dentro de este esquema tradicional?

(ii) Cada definición tipo incluye un *etc.*

(iii) Si dejamos fuera el *etc.* ninguna de las definiciones ofrecidas proporciona condiciones suficientes para la adecuada aplicación de los términos. Algo puede tener *la apariencia* de ser agua o de ser oro, sin serlo. Esto seguirá siendo así aun si ampliamos tales definiciones, mientras estas adicio-

nes satisfagan el requisito de que las definiciones resultantes expresen verdades *a priori* (en el sentido señalado).

(iv) Las definiciones tampoco dan condiciones necesarias de aplicación. Las definiciones ofrecidas tan sólo describen a los objetos en condiciones *normales*, pero éstas pueden cambiar.

3. Contra la tesis tradicional, Putnam ofrece una explicación alternativa acerca de cómo han de "definirse" las palabras para clase. Primero, da cuenta de cómo se *introducen* al lenguaje las palabras para clase. Tal introducción comienza con una referencia a algún conjunto de ejemplares paradigmáticos de la clase en cuestión. Aquí Platts *supondrá* que en esto no hay problemas. En el contexto de esta introducción mantendrá que, por ejemplo, cualquier cosa será agua si es un *líquido de la misma clase* que los que tenemos en el conjunto paradigmático o será oro si es un *metal de la misma clase* de la que ejemplifican los paradigmas. Luego puede eliminarse la referencia a la *mismidad de clase* mediante una investigación científica de las *estructuras explicativas* (que se suponen comunes) en los paradigmas: la estructura molecular o la atómica en los casos del agua y del oro respectivamente. Esta investigación permitirá dar cuenta, luego, de las propiedades "superficiales" comunes y, por ejemplo, podremos decir que algo es oro si tiene la misma estructura atómica que tienen los ejemplares paradigmáticos. Los detalles se esta investigación científica son conocidos, obviamente, *a posteriori*, de tal manera que la verdad de las condiciones de membresía en una clase natural será *a posteriori*, también.

4. El autor supone que el esquema de explicación que ha dado es correcto, aun cuando muy simplificado, para cuando menos algunos casos favorables de palabras para clases naturales relativamente "no teóricas". Algunos comentarios más son pertinentes. El *sentido austero* de una palabra para clases naturales, lo ofrece un enunciado con una forma parecida a la siguiente (más adelante Platts volverá sobre este punto):

(SA) "es oro" es verdadero de α sii α es oro.

Razones de muy diversa índole han impedido que tales *reglas austeras* se acepten como candidatos plausibles de las definiciones *a priori* tradicionales. Algunas de tales razones son: la trivialidad de tales reglas (se dirá más acerca de esto). Razones que se han esgrimido para rechazar tales reglas como posibles descripciones de los significados de las palabras para clases naturales, son que tales reglas no permiten adquirir un dominio de la palabra a quien aún no lo tiene; otra razón es que tales reglas no ofrecen un análisis descomposicional del sentido de las palabras para clases naturales; se ha señalado, además, que tales reglas no pueden servir como guías

en el uso de las palabras, así como tampoco pueden indicar cómo ha de verificarse un enunciado que contenga tales palabras.

Con riesgo de producir confusiones, Platts reune las consideraciones señaladas de la siguiente manera. Por una formulación del *sentido definicional* de una expresión, entenderá una formulación de sus condiciones de aplicación verdaderas, necesarias y suficientes, con cuando menos una de las siguientes propiedades:

(i) es al menos posible que, para algún hablante, el enunciado sirva para proporcionarle un dominio de la expresión en cuestión; *o*

(ii) el enunciado sea "descomposicional" en el sentido de que es al menos posible que dé razón de algunas de las inferencias que comprenden enunciados en los que figura la palabra en cuestión; *o*

(iii) al menos es posible que, para algún hablante, el enunciado sirva como regla guía para el uso de la palabra en cuestión; *o*

(iv) es posible, al menos, que el enunciado sirva como guía acerca de cómo han de verificarse enunciados que contengan la palabra en cuestión.

Al menos dos razones llevan al autor a construir una definición como la anterior: conforme al análisis de Putnam, la referencia de una palabra para clase natural es algo que surge (parcialmente) con anterioridad a su sentido definicional, lo que va en contra de lo que se sostiene conforme a diferentes explicaciones de análisis de significado; por otra parte, la anterior definición pretende reflejar el carácter (*ad hoc*) de las ideas sobre análisis de significado de Putnam, quien en puntos diferentes de sus discusiones parece estar influido por varias de las restricciones posibles anotadas en la definición.

5. Más observaciones preliminares. Dado que, conforme a la explicación de Putnam, la formulación del sentido definicional de una palabra para clases naturales será una verdad (putativa) *a posteriori*, resultado de la investigación empírica de la extensión parcialmente fijada, el hablante ordinario, aun cuando sea competente en el uso de una de tales palabras, no necesita tener idea alguna de la teoría científica mediante la cual se fijaron las condiciones de aplicación, necesarias y suficientes, de aquélla.

Para evitar, entonces, el cargo de que es arbitrario el uso que hace de tales palabras, Putnam propone su "hipótesis de la universalidad de la división de la labor lingüística", conforme a la cual tal uso no es arbitrario en tanto que el mismo siempre puede justificarse, en caso de disputa, apelando a la opinión de los especialistas pertinentes, miembros de la comunidad lingüística del hablante.

Sin embargo, es importante distinguir claramente entre este fenómeno "sociolingüístico" de la división de la labor lingüística y el rasgo "lógico semántico" de las palabras para clases naturales, la prioridad parcial de la referencia frente al sentido definicional.

6. Para señalar cuál sea el significado de una palabra para clases naturales, Putnam ofrece la siguiente "descripción en forma normal" del significado de "agua":

Marcadores sintácticos	*Marcadores semánticos*	*Estereotipo*	*Extensión*
Sustantivo de masa; concreto.	Clase natural; líquido.	Incolora; transparente; insípida; quita la sed; etc.	H_2O (con más o menos impurezas)

Pero, en el mismo artículo en que esto aparece, Putnam nada nos dice acerca de la naturaleza general del significado, del objetivo teórico de la noción misma de significado. Quizás esto sea intencional pero, a pesar de todo, sigue en pie la cuestión acerca de cuáles "hechos" tienen que ver con la noción de significado y de cómo se relacionan con ella.

7. Lo insatisfactorio del tratamiento del significado, por parte de Putnam, surge al considerar sus razones para incluir como "parte del significado" de una palabra para una clase natural lo que él denomina su estereotipo. Esto necesita de un apoyo razonado para eliminar la sospecha de que no es un residuo del fantasma del modelo *a priori* del análisis del significado que Putnam mismo ha criticado de manera tan convincente.

8. Putnam sostiene que para comunicar el uso aproximado de una palabra para una clase natural hay que comunicar ciertos hechos centrales que incluyen la descripción estereotípica.

Putnam, sin embargo, deja inarticuladas y sin defensa las conexiones que se supone se dan entre "comunicar el uso aproximado" y una representación teórica del significado. Parece que apoya una de las ideas tradicionales acerca del análisis del significado, la que señala que cualquier regla que pretenda describir el significado de una expresión debe ser tal que pueda usarse para hacer que quien carece del dominio de la expresión, lo adquiera por su medio. Pero es incoherente imponer esta limitación como un requisito completamente general sobre las "descripciones en forma normal" del significado y Putnam no da razón alguna para retenerlo en el caso especial de las palabras para clases naturales.

Parece que no es posible formular una tesis acerca de lo que *tiene* que hacerse para "comunicar el uso" de una palabra. Pero, por otra parte, la tesis de Putnam no indica que para comunicar tal uso se tenga que *enunciar* expresamente la descripción estereotípica sino que, más bien, la tesis debe ser que tal descripción estereotípica ha de poder extraerla la audiencia de las muestras a las que se aplica la palabra, cuando la descripción no se enuncia expresamente.

Supongamos que es verdad que, para que el auditorio llegue a adquirir incluso una comprensión aproximada de la palabra en cuestión, mediante la mostración de casos paradigmáticos de aplicación, debe formarse algunas creencias pertinentes acerca de la clase que (parcialmente) se muestra; y que la pertinencia de las creencias consiste en que las mismas sean acerca de rasgos típicos de la clase. Sin embargo, de esto no se sigue que haya algún conjunto único de rasgos típicos que la audiencia deba captar.

Bastaría con que las creencias que adquiriera nuestro auditorio fueran inteligibles. Podemos incluso suponer que haya personas con una *concepción* diferente a la nuestra acerca de una determinada clase natural sin que, por esto, tengan un *concepto* diferente de la misma. Aun sin poder dar un detalle de los rasgos típicos de un tigre, por ejemplo, podemos inteligiblemente enseñar el uso de la palabra "tigre".

9. En "The Meaning of 'Meaning' ", Putnam presenta otro argumento en apoyo de la retención del estereotipo. Nos dice: "... La comunicación significativa requiere que la gente sepa algo acerca de lo que está hablando... Lo que sostengo es que se requiere que los hablantes sepan algo acerca de los tigres (estereotípicos) para que se considere que han adquirido la palabra "tigre"...". Pero, de esto, claramente, no se sigue que haya algo único que los hablantes *deban* saber y, además, no es claro por qué ha de requerirse que los hablantes tengan un conocimiento o que éste sea de tigres estereotípicos.

Lo que pretende una teoría de la interpretación es hacer sensata la conducta lingüística de una persona. Para lograr esto es central atribuirle, a la persona en cuestión, creencias y no conocimientos; y estas creencias pueden ser falsas o no, acerca de tigres estereotípicos.

Algo de suma importancia en este enfoque es que el mismo tiene fundamentos totalmente diferentes a los que Putnam considera. Nada tiene que ver con el requisito de que "la gente sepa algo acerca de lo que está hablando". Está en plena relación con consideraciones muy generales acerca del papel que desempeña una teoría de la interpretación, un papel que fija el objetivo teórico de la noción misma de significado.

10. Nuevamente en "The Meaning of 'Meaning'" se sugiere otro argumento. Este comienza con la cuestión acerca de cómo ha de describirse el significado de una palabra para una clase natural.

Putnam rechaza las reglas austeras (cf. § 4) porque nada dicen acerca del significado de la palabra que presentan. Si en su lugar presentamos una regla como

(A_2) "es agua" es verdadero de α sii α es H_2O,

Putnam señala que ésta no es una descripción del *significado* de la palabra "agua". Podríamos explicar este rechazo de Putnam señalando que (A_2) falsificaría la *comprensión* que tiene el hablante ordinario de su lenguaje al falsificar su conocimiento. Lo que se necesita es una regla "informativa" que conecte la comprensión del hablante ordinario con lo que éste *sabe*. Y esto lo satisfaría la especificación estereotípica. Sin embargo, es simplemente falso decir que una regla austera *nada* dice acerca del significado de la palabra que presenta, sino que al pasar de la *designación* de una expresión (al lado izquierdo del bicondicional) al *uso* de la misma, registra, en castellano, un hecho contingente, aprendible acerca del castellano, que es lo que una teoría del significado debería intentar registrar.

Quizás el último argumento que Platts atribuye a Putnam no sea, para éste, un argumento, pues dado que un hablante competente no puede, en general, enunciar la regla que da el sentido definicional de una palabra para una clase natural, ¿hay algo más que deba saber por virtud de tal competencia? Entender una palabra, ¿debe implicar estar en algún estado cognoscitivo ("informativo") determinado? Esta es una petición de principio en el argumento que le atribuimos a Putnam.

Dentro de su estructura teórica, Platts no tiene dificultad alguna para explicar la inaceptabilidad de (A_2). La aceptabilidad de una teoría del significado está determinada, dentro de esta estructura, por el papel que desempeña dentro de una teoría interpretativa total del uso lingüístico que conduce a adscripciones plausibles de actitudes proposicionales, de creencias entre tales actitudes. Por esta razón se debe rechazar (A_2), por lo inadecuado que es dentro de una teoría del significado, dado el papel que tal teoría desempeña dentro de una explicación total del uso lingüístico.

11. Aun cuando se haya eliminado la ilusión de trivialidad total acerca de las reglas austeras, es probable que aún quede alguna ansiedad a este respecto. Consideremos el caso en que intentamos construir una teoría del significado del castellano *en* castellano. Nuestro lenguaje objeto es parte de nuestro metalenguaje, el que excede al primero sólo por contener cierto vocabulario semántico. En tal caso, ¿no facilitaría demasiado nuestra tarea emplear reglas austeras?

Acerca de esto Platts hace tres observaciones breves. Primero, la aceptabilidad de una teoría del significado está determinada por la aceptabilidad de una teoría *interpretativa* total, de la cual es un componente. Al construir una teoría interpretativa del significado, estamos intentando construir una representación teórica (proposicional) de nuestras capacidades prácticas usuales. Ciertamente, cuando dentro de las restricciones de la interpretación en general presentamos una teoría del significado que incluye reglas austeras, aún quedan cuestiones por resolver acerca de *qué* verdades enuncian tales reglas austeras; pero la idea de que *debe* haber alguna forma no austera de decir cuáles son esas verdades, es una idea de naturaleza reductiva que es incoherente si se la adopta como una restricción general.

La siguiente observación explica el uso de Platts de "parecido a" al dar reglas austeras. Aun si, en general, usamos una estructura de interpretación austera en el caso inglés-inglés, surgen cuestiones sustanciales acerca de la "gramática lógica" de las expresiones en el lenguaje que se estudia. Por ejemplo, la discusión reciente de la semántica de los términos de masa (*mass terms*) muestra que no hay nada trivial en la construcción de una teoría del significado para el inglés, incluso cuando se acepta la austeridad en principio.

En la tercera observación, Platts señala que al presentar la explicación de Putnam acerca de los términos para clases naturales, siguió de cerca su presentación esbozando un cuadro histórico de cómo se *introducen* tales términos a un lenguaje. Pero lo que importa es distinguir a los términos para clases naturales de otros tipos de términos para clases, por referencia a su funcionamiento *dentro* de un lenguaje. Determinar cuáles sean los criterios a usar para lograr una clara distinción *semántica* entre términos requerirá tanto de una consideración de los objetivos generales de la interpretación, como del papel que desempeñan, dentro de teorías interpretativas, las asignaciones de propiedades estrictamente semánticas; aceptar una austeridad general no resuelve para nosotros esta cuestión.

12. El autor considera un último argumento para la retención del estereotipo. En "Is Semantics Possible?" Putnam plantea la posibilidad de que los limones se tornen azules y "se queden así". Dice que, "con el tiempo limón llegaría a tener un significado con la siguiente representación:

limón: palabra para clase natural	características asociadas: cáscara *azul*, sabor ácido, etc.

Entonces, 'limón' habría cambiado su significado". Esto es, cambios en la descripción estereotípica producen cambios en el significado, así que el esteretipo debe ser *parte* del significado.

Para obtener esta conclusión se requeriría que *cualquier* cambio en el estereotipo implique un cambio de significado. Pero, ¿es esto así?

Platts indica que va a cambiar de terminología. En lugar de hablar de mismidad o de diferencia de significado, hablará de mismidad o diferencia de *conceptos*. Su pregunta no es "¿tiene limón el mismo significado que tenía hace doscientos años?" sino, más bien, "¿expresa limón, ahora, el mismo concepto que expresaba hace doscientos años?"

Debemos distinguir entre *conceptos* y *concepciones*. Por ejemplo, mi concepción de un limón se descubre viendo qué propiedades de los objetos sirven mejor para explicar mi identificación de objetos dados *como* limones (lo sean o no). En muchos casos habré sostenido explícitamente creencias acerca de cuáles sean tales propiedades; cuando tales creencias sean verdaderas, mi estereotipo de un limón constituirá, en general, parte de mi concepción de lo que es un limón.

Las cuestiones acerca de la mismidad o diferencia de conceptos se resuelven mediante cuestiones acerca de cuál sea la teoría interpretativa correcta. Tales cuestiones, a su vez, están limitadas por la condición de plausibilidad en la adscripción de actitudes proposicionales, entre las que se cuentan las creencias.

Una teoría interpretativa que resuelva "cuestiones conceptuales" debe tener como resultado asignaciones plausibles de actitudes proposicionales. Si no se atribuyen inteligiblemente las actitudes proposicionales que se asignan bajo un esquema de mismidad de concepto, entonces hay diferencia de conceptos. Así que la diferencia, o *mismidad*, de creencias tiene una íntima conexión con cuestiones de mismidad o diferencia de conceptos. Por tanto, no hay ninguna fórmula nítida acerca de las relaciones entre mismidad de concepto y mismidad de concepción, pues no hay ninguna fórmula nítida para la interpretación.

13. Las moralejas generales que hemos de extraer de lo anterior son obvias. Claro está que la "fuerza de los hechos" no debe debilitarse; claro está que si todo lo que tenemos son *teorías* del significado, hay algo terriblemente equivocado en todo esto. Pero la solución no es volver a una charla casual, ni siquiera a una charla filosófica casual, acerca de "significado" al enunciar "cuatro verdades" sobre semántica.

Putnam nos dice en un lugar: "Lo que hasta ahora hemos analizado es el sentido predominante de las palabras para clases naturales (o, más bien, la extensión predominante). Pero las palabras para clases naturales poseen, típicamente, diversos sentidos..." Tal expresión es oscura y enigmática.

Hablando en términos generales, dentro de la tradición fregueana

uno encuentra, por ejemplo, la interesante sugerencia de que una regla que pretenda dar el sentido de una expresión, es una especificación adecuada de tal sentido sólo si el conocimiento de la verdad expresada por tal regla bastase para entender esa expresión.

Dada esta idea relativamente exacta, Putnam no ha dicho nada para mostrar que las palabras para clases naturales "poseen, típicamente, diversos sentidos". Ni, dada esa idea, podría dársele sentido alguno a decir que el "sentido predominante" es la "extensión predominante" puesto que así sólo se confunde una noción ligada al conocimiento de verdades, con una ligada al conocimiento de objetos. Hay *una* noción de sentido en el castellano ordinario, como lo muestra Platts en su última oración; pero el alcance que tiene esa noción y las verdades enunciables mediante ella, sobre nuestras teorías del significado y, así, sobre nuestra noción teórica de significado (o sentido), es un problema sustancial.

14. Una moraleja provisional final es menos obvia. Putnam rechaza la idea de que el "problema central" en semántica es: ¿cómo llegamos a entender una nueva oración? Aparentemente, su tesis es que mucho más importante es la pregunta "¿cómo llegamos a entender una nueva *palabra*?"

Frege sostuvo que preguntar por el significado de una palabra aislada, no en el contexto de una proposición, nos obliga casi a caer en el psicologismo. Este pensamiento de Frege parece estar bien ejemplificado en la explicación de Putnam. Su retención de la descripción estereotípica como parte del significado de una palabra para una clase natural se identifica con el psicologismo al dar razón de la comprensión de los hablantes de tales palabras.

El psicologismo, en la teoría de la comprensión, falsifica la fenomenología del uso del lenguaje; esa falsificación recibe su imagen especular "teórica" en explicaciones psicologistas del sentido. Mientras mayor es nuestra competencia en el uso de un término, menos necesitamos apoyarnos en tesis acerca de sus "criterios" asociados de uso en nuestra aplicación de ese término; lo que permite introducir la posibilidad, excluída por la explicación de Putnam, de que la comprensión de ese término trascienda las diferencias en creencias (estereotípicas) de los hablantes.

Platts subraya que nunca ha comprendido plenamente el fundamento de la conjetura de Frege acerca de cómo estamos "casi obligados" a caer en el psicologismo si consideramos a la palabra como la unidad primaria de significado. Pero aun cuando el psicologismo de Putnam tiene otras fuentes, su sorprendente ejemplificación de la advertencia de Frege muy bien podría hacernos reconsiderar tal conjetura.

[*J. A. Robles*]