SOME PROBLEMS OF SELF-REFERENCE IN JOHN BURIDAN

By A. N. PRIOR

Read 21 November 1962

IN all the periods in which their subject has been in a flourishing condition, logicians have devoted considerable attention to paradoxes involving self-reference. In this preoccupation, it is easy to accuse them of pedantry and even of frivolity; but such accusations are a mistake. Some of the paradoxes of Buridan even have a certain grim relevance to our practical predicaments in this nuclear age, and as it were bring together the Russell who gave us his part of *Principia Mathematica* and the Russell who worries about world peace. But even apart from that, paradoxes about self-reference present exceptions or apparent exceptions to logical generalizations of great persuasiveness, and any logician with a scientific conscience is bound to take them seriously.

The puzzles of this sort which I shall be considering in this lecture come from the eighth chapter of Buridan's *Sophismata*.[1] This treatise, like medieval logical treatises generally, is rather unsystematic by modern standards, and we have to gather what Buridan's leading principles are as we go along; its acuteness lies in its details, and in Buridan's eye for ingenious objections to inadequate solutions. The skill of modern logicians, and indeed of some ancient logicians, in developing not only particular proofs and disproofs but large deductive systems, is something which seems to me wholly admirable; but it is useful, and I think it even makes for the construction of better systems in the end, if this activity is continually interrupted by bouts of the philosophic niggling at which the schoolmen were masters, and Buridan perhaps one of the greatest masters.

Chapter 8 of the *Sophismata* contains some twenty 'sophisms' or debatable sentences or arguments, which fall successively into a few broad groups. There are to begin with one or two

[1] I should like here to thank Mr. Peter Geach for not only drawing my attention to this chapter but sending me a copy of it when I was working on these problems in some isolation in New Zealand.

which involve the notions of logical consequence and possibility; for example, he inquires into the validity of the inference 'No proposition is negative, therefore some proposition is negative'. This hinges on the question whether it is possible that no proposition should be negative, the argument against it being that if the proposition 'No proposition is negative' were ever true it would be false, since it is itself a negative proposition. Then there is a small group about propositions which occur as parts of other propositions, as in 'I say that a man is a donkey'—if a man thus says, not that a man is a donkey, but only *that he says that* a man is a donkey, is he right or wrong? After this group we have Buridan's variations on the ancient paradox of the Liar. He asks whether the proposition 'Every proposition is false' would itself be true or false if enunciated when all *other* propositions were certainly false; and again, whether Socrates and Plato speak truly or falsely if Socrates says 'What Plato says is false', and nothing else, and Plato says 'What Socrates says is false', and nothing else; whether Plato speaks truly or falsely if under the same conditions he says 'What Socrates says is *true*', and says nothing else; whether a man speaks truly or falsely if he says 'There are exactly as many true propositions as false ones' when the only other propositions are two obviously true ones and one obviously false; whether a man utters a falsehood or a truth if he simply says 'I am uttering a falsehood' and says nothing else (this is, of course, the original 'Liar' paradox); and finally, whether the conjunctive proposition 'God exists and some conjunctive proposition is false' is true or false if it is the only conjunctive proposition there is. After this we have a group involving the notions of knowledge, doubt and belief. Suppose we say that a proposition is in doubt with a person if and only if he neither knows that it is true nor knows that it is false. Then we can suppose that the proposition 'Socrates knows that the proposition written on the wall is in doubt with him' is written on a certain wall, and nothing else is written there, and Socrates sees it and wonders whether it is true or false, and knows that he is doing this. Buridan asks whether in this case the proposition on the wall would be true or false.

After some further examples of this last sort, Buridan has a final section in which he considers puzzles arising not with statements but with questions, wishes, promises, &c. He asks, for example, what we are to make of the answer 'No' to the question 'Will you answer this question negatively?' Then he considers a situation in which Plato promises to let people over a certain bridge if and only if the first thing they say to him is true, and to throw them in the river if and only if what they say to him is false, and Socrates says to him, 'You will throw me in the river'. This is a puzzle of some literary interest, since there is a very similar one in *Don Quixote*—Sancho Panza, as governor of an island, is asked to adjudicate in a case where people who cross a certain bridge are hanged if they state their purpose falsely, and let go if they state it truly, and a man announces as *his* purpose that he has come over to be hanged.[1] To finish up with, there are three puzzles about conflicting conditional wishes. For example, Socrates wishes to eat if and only if Plato wishes to eat, but Plato wishes to eat if and only if Socrates does *not* wish to eat. This is where we begin to be reminded of contemporary problems of high politics.

I shall not be discussing this last group of puzzles here, but will concentrate on some of the more elementary ones. And before examining some of Buridan's solutions in detail I want to jump the centuries and, for comparison's sake, briefly survey the treatment that such self-reflexive paradoxes have received in our own time. In the preface to the first edition of Whitehead and Russell's *Principia Mathematica* some seven 'contradictions' of this sort are listed, and there is said to be in all of them 'a common characteristic, which we may describe as self-reference or reflexiveness'.[2] F. P. Ramsey, in his 1925 paper on *The Foundations of Mathematics*, divided these into two sharply demarcated groups, of which the first 'consists of contradictions which, were no provision made against them, would occur in a logical or mathematical system itself', while those in the second group 'all contain some reference to thought, language or symbolism'.[3] Typical of the first group is Russell's paradox of the class of all classes which are not members of themselves—this class being, on the face of it, a member of itself if it is not, and not a member of itself if it is. Typical of the second group is the paradox of the Liar—the man who says 'What I am saying is false', and says nothing else, his statement being, on the face of it, true if it is false and false if it is true.

The 'provision' which Ramsey had in mind for preventing

[1] Miguel de Cervantes, *Adventures of Don Quixote de la Mancha*, ch. li (cited in Alonzo Church, *Introduction to Mathematical Logic*, Exercise 15. 10).

[2] Alfred North Whitehead and Bertrand Russell, *Principia Mathematica*, 1st ed., vol. i (1910), pp. 63 ff.

[3] F. P. Ramsey, *The Foundations of Mathematics and Other Logical Essays* (Kegan Paul, 1931), p. 20.

our logic and mathematics from being disfigured by paradoxes of the first sort, was the so-called 'simple theory of types'. Here the paradoxes are resolved by denying that classes of classes are classes at all in the sense in which classes of individuals are. This view is usually associated with the view that classes are in any case 'logical fictions', talk about classes being only an oblique but often handy way of talking about individuals. Thus understood, the theory of types is basically a theory of 'syntactical categories' or, in a broad sense, 'parts of speech'. To say that *x* is a member of the class of smokers is not really to relate two objects, *x* and the class of smokers, but is simply to say, approximately, that *x* smokes, where 'smokes' is not a name but a verb, that is an expression which forms sentences out of names. To say (no doubt falsely) that the class of smokers is a member of the class of 6-membered classes is again not to relate two objects, or even any objects, but simply to say, approximately, that exactly six individuals smoke, where the prefix 'Exactly six individuals' is neither a name nor a verb but a higher-type expression, a numerical quantifier in fact, which forms sentences out of verbs. Talk of classes being members of themselves then expands to a form in which we make some verb its own subject—a form with a bit in it like 'smokes smokes'—which just does not construe, or as the logicians say is 'ill-formed'.

This may seem an awful lot of grammar for dealing with paradoxes of Ramsey's *first* type, but remember that what is being straightened out here is our talk about individuals and classes and numbers; what has to be straightened out to get rid of the other group of paradoxes is our talk about talk. And their solution is usually understood to require not merely a hierarchy of syntactical categories or parts of speech, but a hierarchy of languages. Sentences which are true or false are always true or false in some language, and *that* they are true or false is not itself true or false in that language but in some higher one. These might not be different languages in any ordinary sense, but rather different levels or stages of a single one, but the point remains that an assertion *about* the truth or falsehood of a sentence cannot itself be true or false in the language or level or stage of a language to which the sentence itself belongs. For this reason no sentence can directly or indirectly assert its own falsehood, or for that matter its own truth. This hierarchy of languages is a very fundamental conception in, for example, Tarski's well-known monograph on Truth.[1]

[1] Alfred Tarski, 'The Concept of Truth in Formalised Languages'. Item VIII in *Logic, Semantics and Metamathematics* (Clarendon Press, 1956).

Even before Ramsey wrote his paper there were known to be alternatives to the theory of types as a method of handling the strictly mathematical or logical paradoxes, and some of these alternatives have been very fully developed since. In particular it is perfectly possible to treat classes of individuals, classes of classes, &c., as nameable objects in exactly the same sense as individuals are, provided that we do not over-simplify the relations between objects of these various sorts. One may, for instance, refuse to equate simply φ-ing with being a member of the class of φ-ers, and one may hold, e.g. that the class of classes which are not members of themselves is-not-a-member-of-itself and yet is not a member of the *class* of classes-which-are-not-members-of-themselves. Zermelo's set theory and the modifications of it which have been made by von Neumann and Quine are well-known systems of this broad type.[1]

On the side of the puzzles which Ramsey has classified as linguistic rather than logical or mathematical, the most important development since his time has been the clear demonstration by Gödel, Carnap and others that there is a great deal that *can* be said within a given language or language-level about that language or language-level itself, e.g. we can talk within a language about that language's grammatical structure. What is still generally disallowed, in order to eliminate contradictions, is talk within a language about its relation to the rest of the world, and in particular about questions of the meaning and truth of expressions within it. Taking one of Buridan's examples, a modern logician would insist that the proposition that the proposition 'No proposition is negative' is *true* must belong to a higher language-stratum than that proposition itself, but the proposition that it is *negative* is one that *can* be framed within the very language-stratum in which the proposition itself is framed.

There has also been considerable interest in recent years in what are called 'pragmatic' paradoxes, involving personal attitudes like belief and knowledge. Particular attention has been given, under this head, to a puzzle which appears in the literature in a variety of dramatic guises, e.g. the story is told of a prisoner who is sentenced to be hanged on some one of a number of days, but who is told when he is sentenced that he will not know which day it is until the time comes. He works out that it cannot be the last day because then, all the other days having

[1] See, e.g. W. V. Quine, *From a Logical Point of View* (Harvard University Press, 1953), Item V.

passed, he would know beforehand that it was going to happen that day. He then successively eliminates the other days in the same way, but is nevertheless hanged on one of them, and unexpectedly too. There is no general agreement as to how paradoxes of this sort are to be solved or classified.

It seems to me useful to bear in mind these broad features of current work when examining Buridan's treatment of the same and allied topics, but we would be wise *not* to take it for granted that we know all the answers better than he did; and with a little open-mindedness I think we can find in him not only the material for new formal exercises but also suggestions towards new solutions of our problems. And I shall devote a good deal of time to simply *discussing* Buridan's theories as if he were present and one of us—on the general principle that there is much less to be learnt from the history of philosophy as history than there is from the people we meet in it.

One of the most striking contrasts between Buridan's discussion and, for example, Russell's is the complete absence from Buridan's chapter of any puzzles falling into Ramsey's first or logico-mathematical group. In an earlier chapter of the same work there is indeed something that could be construed in this way. This is the fourth *sophisma* in Chapter 3, to the effect that there is a wider *genus* than the widest one, for the term *genus* itself covers both the widest *genus* and all the less wide ones besides. A distinction between classes of individuals and classes of classes seems clearly called for here, but Buridan's own solution is very brief and not very satisfactory—he seems to think that the secret lies in the distinction between the *word* 'genus' and what the word signifies. Ramsey could have put him right here.

In the eighth chapter, all the puzzles considered do fall into Ramsey's second or linguistic group, apart perhaps from the ones near the end about doubting, &c., which modern writers would classify as 'pragmatic'. It is significant here that Buridan almost invariably uses the term *propositio* to mean simply a bit of language, a spoken or written sentence—a particular noise or inscription. And he never forgets that the very existence of 'propositions' in this sense is a contingent matter, and speaks quite freely of the *annihilation* of propositions (asking what would be the case, for example, if all negative propositions were annihilated). There is just no trace in him of the use of the term 'proposition' to mean, not a sentence, but a supposed abstract entity of which the sentence, or the corresponding 'that' clause, is a name. Nor does he appear to believe that there *are* entities of this sort. There is, indeed, a passage in the discussion of his seventh *sophisma* in which he might seem to hold that *true* sentences do name such abstract entities but *false* sentences do not, and he even at this one point uses the word *propositio* for what does not exist when a sentence is false. He makes in this passage the strange remark *hominem esse asinum nihil est*, and the context makes it clear that he would have equally said *hominem esse animal aliquid est*. It is quite clear from what he says elsewhere, however, particularly under his sixth *sophisma*, that he would *not* have meant by this that there is an abstract entity of which the expression 'that a man is an animal' is a name. The *aliquid* that the expression *hominem esse animal* and even the sentence *homo est animal* stands for is not an abstract entity but simply a *man*, a man being an animal, and *hominem esse asinum* or *homo est asinus* in this sense stands for nothing because there is no such object as a man being a donkey.

What is really illuminating in Buridan, however, is not this rather curious material about what sentences 'stand for' but his treatment of the question as to what they *mean* or signify. What he does is not so much to answer this question as to transpose it. The transposition comes out most clearly in his account of what it is for a sentence to be true. A sentence is true if and only if *sicut significat, ita est*, or as he sometimes says if *qualiter significat, ita est*; this, though, is only a first approximation—the final version is: *qualitercumque significat, ita est*. This is not easy to put into English, but the important point is that Buridan does not say that a sentence is true if and only if *what* it signifies, or *whatever* it signifies, is so or is the case; what he says is rather that a sentence is true if and only if *however* it signifies that things are, *thus* they are. He gets rid of the suggestion of *objects* that are signified by sentences by beginning his definition not with a generalized noun but with a generalized adverb, *qualitercumque*.

What this means, in modern terms, is that the hierarchy of parts of speech is relevant not only to our talk about classes and the like, but also to our talk about meaning. Modern writers have seen this also; for example, Russell in his 1924 paper on 'Logical Atomism' says this: 'When two words have meanings of different logical types, the relations of the words to what they mean are of different types; that is to say, there is not one relation of meaning between words and what they stand for, but as many relations of meaning, each of a different logical type, as there are logical types among the objects for which they are words.'[1] By current standards even this, with its easy talk of 'objects' of different logical types and of 'relations' of meaning which hold in each case, is extremely loose and a little misleading, but I think we can put what Russell is getting at in the following way: In spite of some recent objections, I think it can be argued that what a proper name means is simply the object that it names—'Fido', the name, means or names Fido the dog. But a verb cannot have this sort of meaning, for verbs do not name anything, and if we wrote, for example, ' "Runs" means runs' this would just be a senseless sentence, with two verbs and only one subject. In English and other languages we invent abstract nouns to meet this difficulty, and we could say that 'runs' means or signifies the activity of running, or more shortly that 'runs' means running. But we misunderstand the function of abstract nouns if we think that there is an object called 'running' which the verb 'runs' names, or even does something else to ('connotes' it or what have you). The fact of the matter is that here the word 'means' has no meaning in itself but is just part of the expression 'means (blank)-*ing*', which constructs a sentence not out of two names but out of a name and a verb. When we come to the meaning of *sentences*, the word 'means' is again without meaning on its own, but is part of the expression 'means *that*', which constructs a sentence not out of two names but out of a name and a sentence—'A man is a donkey' *means that* a man is a donkey. So we need not ask what is named by the clause 'that a man is a donkey'; the word 'that' does not belong here but with the 'means' that precedes it, and what is left, 'a man is a donkey', names nothing because it is not a name but a (subordinate) sentence.

After this excursion into grammar, we can re-state Buridan's definition of truth as follows: A sentence x is a true one if and only if for any p, if x means that p, then it is the case that p; or more shortly, if for any p, if x means that p, then p. This is much simpler than any of Tarski's definitions of truth for the various languages that he considers. It ought in fairness to be added that part of Tarski's aim was to avoid the use of 'intensional' conceptions like that of 'meaning'; but it is certainly worth noting that if we do not restrict ourselves in this way, and get our grammar straight, it *is* possible to define 'true' very straightforwardly.

Does Buridan's definition, however, avoid the necessity of a language-hierarchy? On the face of it, it does not, but merely

[1] Bertrand Russell, *Logic and Knowledge*, pp. 332–3.

traces the systematic ambiguity of 'true' to a similar ambiguity in the more basic conception of 'meaning'. And I am not now referring to the different syntactical types involved in the meaning of different types of expression, but to the ambiguity which still seems to remain when we confine our attention to the meaning of *sentences*, i.e. to the sort of meaning that is always a meaning *that* something-or-other. For suppose I utter the sentence 'The sentence I am uttering is false', and utter no other sentence; then it would appear that what this sentence means, and all that it means, is *that* the sentence I am uttering is false, and this is therefore true, by the definition of 'truth' just given, if and only if the sentence I am uttering *is* false, and we are back with the 'Liar' paradox. It therefore seems necessary to say that 'meaning that *p*' must always be 'meaning in a language L that *p*'; and *that* a sentence means in a language L that *p*, cannot itself be said by a sentence of the language L.

Buridan, as we have seen, was as familiar with the 'Liar' paradox, and variations upon it, as anyone has ever been, and indeed it is precisely in this context that his definition of truth appears. Language-hierarchies of the systematic type that we meet with in such writers as Tarski are a comparatively modern invention, but Buridan does consider and reject certain solutions which it would be easy to put into the language-hierarchy form. For example, when considering the proposition 'Every proposition is false', supposed to have been put forward by Socrates after all other propositions had been annihilated but false ones, he says that we might understand the proposition of Socrates simply as a comment on everything that was being said in the time just preceding its own appearance on the scene, and then it would be quite straightforwardly true. But, he very properly asks, what happens if we *don't* understand it in this way, but understand it as referring to all propositions in being *at the time*, itself included? He tells us that according to some it just cannot be so understood, because in a proposition which contains terms which themselves stand for propositions, these terms cannot stand for that proposition itself but only for all others. This, however, Buridan says, won't do, for *quod aliquis intelligit, de illo potest loqui*, and as it is certainly possible for someone to think about all propositions whatsoever (past, present, and to come), what he thinks about them can be expressed in a proposition, which will inevitably be itself among those intended.

What else, then, can we do about these paradoxes? Buridan mentions, but rejects out of hand, the solution that some propositions can after all be true and false at once. He then mentions another, which he says that he himself formerly thought satisfactory, to the effect that every proposition, whatever else it may signify or assert, signifies or asserts, by its very form as a proposition, that it is itself true. Any proposition, therefore, which asserts or implies its own falsehood asserts both its falsehood and its truth, and is bound to be in fact false, since at least *something* that it asserts to be the case is not so. We cannot pass back from its falsehood, thus established, to the conclusion that it is after all true, since it says that it is false and things are as it says they are; for things are not *entirely* as it says they are, part of what it says being that it is true.

To this former view of his own, Buridan now objects that propositions do *not* in general signify in virtue of their very form that they are themselves true, because if you take, say, the proposition 'A man is an animal', its terms are 'of first intention', i.e. non-linguistic, while the proposition 'The proposition "A man is an animal" is true' contains terms 'of second intention', i.e. terms referring to pieces of language. Buridan is led by these considerations to distinguish between what a proposition 'formally' signifies and what it 'virtually' signifies. What it 'virtually' signifies is what follows from the proposition itself together with a proposition correctly describing the circumstances of its utterance. In particular, from a proposition x together with the proposition 'The proposition x exists' we may infer the proposition 'The proposition x is true'. And a proposition is only true if all that it signifies, formally *or* virtually, is so. From this point on the argument is very much as before, but I shall not follow it out in detail because this later position of Buridan's seems open to a quite fundamental objection. Since he employs this term 'formal signification' in such a sense that 'What I am now saying is false', for example, formally signifies that what I am now saying is false, and nothing else, we can re-state the paradox in terms of formal signification without bringing in truth and falsehood at all. We simply suppose a person to say 'What is formally signified by this sentence is not the case', and ask about this sentence, not whether it is true or false, but whether things are or are not as it formally signifies that they are, and the answer is that they are if they aren't and they aren't if they are.

This is a transformation of the paradox which suggests itself more readily to a modern logician, accustomed to the use of abstract symbolism, than it would to a medieval one. Suppose I write 'It is φ that p' for *any* sentence formed from the sentence '*p*', for example 'It is not the case that *p*', 'It is possible that *p*', 'It is signified by the sentence *x* that *p*', 'It is feared by the person *y* that *p*', &c., I can then construct the following formula:

> It is φ that, for any *p*, if it is φ that *p*, then it is not the case that *p*; and for no other *p* is it φ that *p*.

From anything of this form it is possible to deduce contradictory consequences, by quite elementary logical processes. There can therefore be no φ which will turn a complex of this form into a true sentence. For example, none of the following can possibly be true:

1. It is being brought about by James that whatever is being brought about by James is not the case; and nothing but this is being brought about by James.
2. It is feared by James that nothing that is feared by James is the case; and nothing but this is feared by James.
3. It is apparently feared by James that nothing that is apparently feared by James is the case; and nothing but this is apparently feared by James.
4. It is signified by *x* that nothing that is signified by *x* is the case; and nothing but this is signified by *x*.
5. It is signified by *x* (so far as *x* signifies anything at all) that whatever is signified by *x* (so far as *x* signifies anything at all) is not the case; and there is nothing else that is (signified by *x*, so far as *x* signifies anything at all).
6. It is conventionally (normally, formally) signified by *x* that nothing that is conventionally (normally, formally) signified by *x* is the case; and nothing that is conventionally (normally, formally) signified by *x*.

At least, none of these is true if the expression substituted for 'φ' is used in the same way throughout the sentence. (If, for example, 'signified' means 'signified in L' in one occurrence and 'signified in M' in another, it is quite a different story.) And any semantics which is to avoid inconsistency must have some means of blocking the introduction of φ's with which sentences of this general form are constructible and provable. Buridan's later theory, so far as I can see, fails to meet this requirement.

But what of his *earlier* theory? This it seems to me, *is* logically workable; it is, at all events, not immediately open to the above objection, since it involves no sense of 'it is signified by *x* that' for which a sentence of the above form would be provable. Its method of preventing this has its own repercussions, some of them perhaps not too palatable; but I fear we must reconcile ourselves to the fact that, however it is conducted, Semantics is a mess. (A theologian of some logical competence once described this as 'a sign of our creaturely status', but even God's language, if such there be, and if it is consistent, must be subject to the same limitations.) It is of some interest that the solution now proposed (the younger Buridan's) has been defended in our own period by a very great logician indeed, namely Charles Sanders Peirce.[1] Peirce did not attempt a detailed formalization of the position, and I have myself only looked at the beginnings of such a development, but I am fairly confident it can be done. In other words, a language *can* contain its own semantics, that is to say its own theory of meaning, provided that this semantics contains the law that for any sentence *x*, *x* means that *x* is true. To set the whole thing out in a fully formalized way, we would need to introduce a symbol, say '*M*', for 'means that', and write '*Mxp*' for '*x* means that *p*', and then with this and ordinary logical symbols we could formulate, and assert as a law of the system, the sentence '*x* means that for all *p*, if *x* means that *p*, then *p*', i.e. '*x* signifies that, however *x* signifies that things are, thus they are', or '*x* means that *x* is true'. This law is not intended as a definition of 'means that', and as such would be circular and absurd, nor does it assert that *all* that a sentence *x* means is that *x* is true, but it does say that any sentence *x* means this, whatever else it may mean besides. The Liar paradox could then be disposed of exactly as the younger Buridan did dispose of it.

What I am really suggesting now is that the fault of Buridan's later theory, and the source of its inconsistency, is just its half-heartedness. When Buridan, in objecting to his own earlier theory, makes so much of the distinction between sentences which do and sentences which do not contain terms of second intention, he has already sold the pass to the proponents of language hierarchies; a man who is really determined to abandon these—a whole-hearted Presbyterian in semantics, as we might say—will not attach much weight to such arguments, drawn as they are from the armoury of linguistic Prelacy. The fact seems to be that if any sentence *could* be about the semantics of its own language, then all sentences of that language to a certain extent *must* be about its semantics, though in general they will be about other things as well.

Let us not, however, be over-violent here, and replace lin-

[1] *Collected Papers of C. S. Peirce*, 5. 340.

guistic feudalism by a new totalitarianism. We must live and let live. It would be foolish to deny that the word 'means' must be relativized to a language—words don't just 'mean' on their own; 'meaning' *is* always 'meaning in a language L'. And the language L in question *could* be one which does not itself contain the expression 'means in L', and which thereby gains various simplicities. But it could also be one which does contain this expression; for this a price must be paid, but it can be a price less than inconsistency.

I would envisage such a 'Buridanian' language as having a syntax of a broadly Russellian type, with a sharp distinction made between genuine proper names and definite descriptions. In particular, the enclosing of a sentence in quotation-marks should not be thought of as forming a genuine proper name of that sentence, but rather as an abbreviated description. ('The sentence "Grass is green" ' would abridge something like 'The sentence formed by writing a Gee followed by an Ar followed by an Ay', &c., &c.) A genuine proper name would have no internal logical structure. But a 'Buridanian' language *would* contain genuine proper names of its own expressions; in fact, in the law '*x* means that *x* is true', the variable must be thought of as one keeping a place for precisely such a proper name. (The law might, incidentally, have to be enunciated in the qualified form 'If *x* means-that anything'—i.e. 'For any *p*, if *x* means that *p*'—'*x* means that *x* is true'; since a proper name is not by its very form a name of a sentence rather than of some other object.) Such a proper name could, moreover, be a name of a sentence in which this name itself occurs, e.g. 'A' could be a name of the sentence 'A is false'. If A were not a genuine proper name but just an abbreviation of the description 'The sentence "A is false" ' this would not be possible, as we would never be able to give the fully expanded form of this description; but a mere proper name would not require any such expansion.

Buridan himself frequently used letters as proper names of sentences, and described them as precisely that. He made it clear, and a modern refurbishing of his semantics would also have to make it clear, that these are proper names not of sentence 'types' but of particular utterances and inscriptions. If 'A', for example, is the proper name of the following inscription: *A is false*, then it is not the name of the following exactly similar (or as is often said 'equiform') but numerically different inscription: *A is false*. And if it is in *this* sense of 'sentence' that all sentences signify their own truth, it follows that even a pair of equiform sentences are never *quite* synonymous. '2 and 2 are 4', for instance, means that 2 and 2 are 4 and that *that* inscription back there is a true sentence; while the following: '2 and 2 are 4', means that 2 and 2 are 4 (this much meaning the two inscriptions have in common) and that *this other* inscription (the nearest one to here) is a true sentence. Further, two equiform inscriptions may not always even have the same truth-value (a consequence which Buridan quite boldly drew). For example, if 'A' is the proper name of this inscription: *A is false*, then A in fact *is* false, but precisely because of this the following inscription: *A is false*, is true. For the first inscription asserts its own falsehood (and, of course, like all inscriptions, its own truth), but the second inscription asserts, not *its own* falsehood, but the falsehood of the *first* inscription (together with *its* own truth; but now there is no contradiction, only a difference in truth-value between equiform inscriptions).

So-called structural-descriptive names of sentences (like 'The sentence formed by writing a Gee followed by an Ar', &c., &c.) will in consequence not be even genuine *descriptions* of sentences, in the sense of 'sentence' intended, but will refer rather to classes of equiform sentences; and the rules which give the meaning of particular sentences will be somewhat complicated. In many cases, nevertheless, equiform inscriptions *will* have the same truth-value, e.g. all inscriptions equiform with this one: *No proposition is negative*, are false; and their differences in meaning can in many contexts be ignored, so that no harm is done by talking about, say, 'The sentence "No proposition is negative" ' when what is really intended is 'All sentences equiform with the sentence "No proposition is negative" '. (And I shall myself indulge in this harmless laxity below.)

With this particular example, we do run into a difficulty, though not an insuperable one, in connexion with the problem which we have already found Buridan raising about it. Buridan insists that the proposition 'No proposition is negative' must be classified as a 'possible' one because things could be as it signifies, even though it could not possibly be true. It cannot possibly be true because it will only be true if it exists, and if it exists there will be at least one negative proposition, namely itself. But if God were to annihilate all negative propositions there would in fact be no negative propositions, even if this were not then being *asserted* in any proposition at all. In short, *it can be that no proposition is negative, though it cannot be that 'No proposition is negative' is true.* Up to this point Buridan's reasoning seems to me quite conclusive and extremely important. Numerous modern writers have insisted that 'possibility' is in the first instance a property of sentences; there are, they say, no possibilities in things themselves, which are simply so or not so; and to say that some state of affairs is possible (that is, to say with respect to some p that it is possible that p) is just to say that the sentence which expresses this state of affairs (that is, the sentence x such that x means that p) has some property or other. What this property is supposed to be is a little obscure, but Buridan's example at least shows that it cannot plausibly be possible-truth. Buridan is still prepared, nevertheless, to use 'possible' as an adjective predicable of sentences, and attempts what one would have thought to be a more hopeful task than the converse modern one, the definition of this 'possibility' of sentences in terms of the possibility of states of affairs (not *vice versa*). A sentence x, he says in effect, is possible not only if it could be true, but also (even when it *couldn't* be true) if things could be as x says they are, or in modern formal terms, if for all p, if x means that p, then it could be that p.

If, however, we adopt the semantics of the younger Buridan, this account of 'possibility' in sentences won't quite do. For according to this semantics, one thing that is meant by any sentence x is precisely that x is true; and in particular, one thing that is meant by the proposition 'No proposition is negative' is precisely that the proposition 'No proposition is negative' is true. If we adopt Buridan's later distinction between 'formal' and 'virtual' signification, we can escape this difficulty by saying that a sentence is possible if everything that is *formally* signified by it could be the case, and since the proposition 'No proposition is negative' formally signifies only that no proposition is negative, and *not* that the proposition 'No proposition is negative' is true, we can classify this proposition as 'possible' because it could be that no proposition is negative, even if the other thing could not be. But we have already seen where this notion of 'formal signification' leads us—either back to the paradox of the Liar, or, as the only means of escaping this, back into the Babylonish captivity of a hierarchy of languages.

There are, however, ways out of this predicament which are quite simple and I think quite satisfactory. In the first place, I don't see that there *has* to be a sense of 'possible', as an adjective predicable of sentences, which is distinct from 'possibly true'. We can still say that it could be the case that no proposition is negative, *without* saying that the proposition 'No proposition is negative' is thereby classifiable as a 'possible proposition'. But if we do insist on using this language, we can define 'possible' as applied to sentences in a more indirect way, namely by saying that a sentence *x* goes into the 'possible' class if and only if the sentence formed by prefixing 'It could be that' to this sentence *x* is a true one. This does give the distinction that Buridan wanted to make. The proposition 'No proposition is negative' is 'possible' in the sense that the proposition 'It could be that no proposition is negative' is true. This longer proposition, it must be admitted, signifies not only that it could be that no proposition is negative but also that the proposition 'It could be that no proposition is negative' is true; but to say that the proposition with 'It could be' in it *is* true is a different thing from saying that the proposition without that addition *could be* true; so Buridan's distinction is still preserved.

The semantics which I have sketched might prove to be in its details (despite the Russellian character of the associated syntax) not unlike the Zermelo-Quine alternative to the theory of types, and I cannot help feeling that it is much more called for. For the simple theory of types, especially in the forms in which it is now propounded by Polish logicians such as Suszko and Borkowski, seem to me not at all burdensome, and anyway even the Zermelo-Quine logic itself has to have *some* distinctions of syntactical categories—a name, for example, is still something different from a sentence. The only gain which this logic brings is a rather technical one, a limitation of the kinds of variables that need to be bound by quantifiers, and I don't think even this advantage can be plausibly carried over into non-mathematical contexts. But a hierarchy of languages, as opposed to a hierarchy of parts of speech, really *is* a lot for us to have to carry around, and if the theories of the younger Buridan promise a way out of it, they are certainly worth looking into.