What then can transpire between the first and second days to make trust in the teacher less reasonable? Given the idealizations we have assumed, the only relevant difference between the first and second days is that by the second day it will be known whether or not an examination occurred on the first. If an examination has been given on the first day, then the teacher's trustworthiness will be, if anything, enhanced, for his announcement will have come true, the students having had no reason to expect it then. But if an examination has not been given, then the teacher's trustworthiness will indeed be placed in question, and the students' caution confirmed; but in a curious way. It will not be because the teacher will have been proved a liar; a surprise examination may still be given. The reason is rather that events will have proved him to be a purveyor, by implication, of incredible propositions. Such people cannot be trusted. You cannot trust a man if he tells you things that you cannot believe.

This comes about because, if there is no examination on the first day, then the teacher's announcement becomes in effect the one-day paradoxical announcement considered above. For if we add $\sim e_1$ to what the teacher has said we can deduce $e_2 \cdot \sim J_2 e_2$, something that the students can't believe on the second day.

Now the students must envisage this paradoxical second-day test of their teacher's trustworthiness if their first-day doubt about their continued belief in his announcement is to be reasonably combined with their first-day trusting belief in it. But as we have seen, this doubt must be present if they are to believe it at all.

So either directly with A5b or indirectly without it, the surprise-examination paradox reduces to the phenomenon of incredible though possibly true propositions, and it should redound to the credit of modal logic that it helps us to see this.

ROBERT BINKLEY

University of Western Ontario

PRESUPPOSITION, IMPLICATION, AND SELF-REFERENCE *

THE two aims of this paper are, first, to explicate the semantic relation of *presupposition* among sentences, and, second, to employ the distinctions made in this explication in a discussion of certain paradoxes of self-reference. Section I will explore informally the distinction between presupposition and im-

* An earlier version of this paper was read at Duke University on May 12, 1967, and sections I and II were included in a paper presented at a symposium on free logic held at Michigan State University on June 9 and 10, 1967. Acknowledgments and bibliographical references have been collected in a note at the end.

plication. In section II we construct a formal paradigm in which this distinction corresponds to a substantial difference. Section III explores the relations between presupposition and truth, and, in section IV, the *Liar* and some similar paradoxes are finally broached.

I

The best known source for the concept of presupposition is the view (a.o. Strawson's) that a property cannot be either truly or falsely attributed to what does not exist. Thus, the sentence "The King of France (in 1967) is bald" is neither true nor false, on this view, *because* the King of France does not exist.[1] The explicit characterization of *presupposes* is therefore given by

1. *A presupposes B* if and only if *A* is neither true nor false unless *B* is true.

This is equivalent to

2. *A* presupposes *B* if and only if
   (a) if *A* is true then *B* is true,
   (b) if *A* is false then *B* is true.

From this equivalence it is clear that presupposition is a trivial semantic relation if we hold to the principle of bivalence (that every sentence is, in any possible situation, either true or false). In that case, every sentence presupposes all and only the universally valid sentences.

Why can we not say that, on Strawson's position, "The King of France is bald" *implies* "The King of France exists"? The reason is that presupposition differs in two main ways from implication, on any accepted account of the latter relation. For implication, *modus ponens* is accepted as valid; from 2(a) it is clear that an analogue of *modus ponens* holds also for presupposition. But *modus tollens* is also generally accepted as valid for implication; what of its analogue for presupposition? This question cannot be answered unless we first settle on the meaning of negation. We shall understand the negation of a sentence *A* to be true (respectively, false) if and only if *A* is false (respectively, true). This is not the only possible convention, but it has the virtue of yielding the convenient further characterization of presupposition:

3. *A* presupposes *B* if and only if:
   (a) if *A* is true then *B* is true,
   (b) if (*not-A*) is true then *B* is true.

[1] The question whether a sentence is true may be said to make sense only relative to an interpretation, or to what is intended by it, and so on. Except when dealing with artificial languages, we shall avoid this issue by assuming that there is a "correct" interpretation known to the reader.

From 3 it is clear that the analogue to *modus tollens*

4. *A* presupposes *B*
   (not-B)
   Therefore, (not-*A*)

is not valid; if the premises are true, the conclusion is not true (and not false). Secondly, 3(b) shows that the argument

5. *A* presupposes *B*
   (not-A)
   Therefore, *B*

is valid: if its premises are true, so is its conclusion. But of course the analogue of 5 for implication does not hold.

Thus presupposition and implication are not the same, but they have something in common. What they have in common is that, if *A* either presupposes or implies *B*, the argument from *A* to *B* is valid. This is itself a semantic relation, which we shall here call "necessitation":

6. *A necessitates B* if and only if, whenever *A* is true, *B* is also true.[2]

We see here an obvious possible cause of confusion between implication and presupposition. The standard logic text generally begins with an explanation that an argument is valid if and only if the conjunction of the premises implies the conclusion, if and only if the corresponding conditional is logically true. This explanation then justifies the text's concern with validating arguments (respectively, with proving logical truths) alone. But the standard logic text is concerned only with a case for which the principle of bivalence holds, and we must resist the temptation to extrapolate its teachings to other contexts.

From 3 and 6 we obtain finally the equivalence

7. *A* presupposes *B* if and only if:
   (a) *A* necessitates *B*,
   (b) (*not-A*) necessitates *B*.

Thus we have an explication of *presupposition* in terms of standard semantic notions.

Before turning to the task of making our account precise, we may briefly consider two questions. The first notes that we use 'if . . . then' in, for example, 3 and 6; how is this conditional to be understood? For the sake of clarity, I propose that this be understood as the ma-

[2] This relationship is usually called "semantic entailment," but in the present context that terminology might be confusing. Note that the 'whenever' does not refer to times, but to possible situations.

terial conditional. This yields at once the limiting case that, if *B* is universally valid, every sentence presuppose it. The second question notes that every case of implication or presupposition is a case of necessitation; does the converse hold? This question is not necessarily decidable at this point, because we have not given a complete account of implication (nor presume to be able to do so). But in section III we present a case of necessitation which is not one of presupposition and which, we shall argue, is also not a case of implication.

II

There are two reasons for constructing a formal paradigm at this point. The first is to show the possibility of having distinct semantic relations of implication, necessitation, and presupposition. That is, we wish to show that the distinctions we have drawn are not merely verbal distinctions, but are capable of expressing nonequivalent concepts. The second reason is an obvious one: the notions of implication and presupposition play an important role in certain philosophical arguments. A case in point is the derivation of paradoxes such as the *Liar,* which we shall examine below. A formal paradigm will give us a means of testing such reasoning.

This may suggest that we mean to construct an axiom system. This is not so. Instead, we intend to serve the two purposes indicated by outlining the construction of a certain kind of artificial language. This kind of language is meant to provide a model for the kind of discourse in which presuppositions are important. Because of our explication of presupposition in terms of necessitation, we shall be able to concentrate on the relations of implication and necessitation alone. Our construction will have to be such that: a case of necessitation need not be a case of implication or of presupposition; a case of presupposition need not be a case of implication; and sentences of which some presupposition is not true, are neither true nor false.

An artificial language has two parts: a *syntax* and a *semantics*. The syntax comprises a *vocabulary* and a *grammar,* which together generate the set of its *sentences*. The semantics defines what I shall call the admissible *valuations*; that is, it delineates the possible ways in which these sentences could be true or false together. (In general, the class of admissible valuations is defined in a rather roundabout way, through a (partial) interpretation of the vocabulary.)

The languages whose construction we shall now outline, we call *presuppositional languages*. By this I mean that in their semantics we make it possible for presuppositions to be made explicit, and we countenance the possibility that, for some of the sentences, presuppositions may fail.

We use 'L' to refer to an arbitrary language of this kind and describe those features which make it a presuppositional language. First, we require that the vocabulary include a negation sign and a disjunction sign (other propositional connectives may be defined in terms of these). I may as well say at once that I will not be satisfied if negation and disjunction do not obey the laws of classical logic. With respect to logic I am conservative: I would resist any imperialism on behalf of classical logic; I would not accept the idea that it is applicable to all contexts or that it is sufficient for all (important) purposes, but on the other hand I have no inclination to change it.

Secondly, the vocabulary may contain a special sign for implication. The material conditional $(A \supset B)$ may be defined as $(\sim A \vee B)$. But some other sign may be specified as *the* implication sign. Here we shall require that the semantic assertion that A implies B be so construed that, at least, it is true only if $(A \supset B)$ is logically true. That is, logical truth of the material conditional is a provisionally acceptable explication of implication, and so is any account under which implication entails this logical truth.

In the semantics, we need two preliminary notions. First, the semantics must specify a relation N of necessitation ("nonclassical necessitation") to cover those cases which cannot be cases of implication. How this relation N is specified in the semantics need not concern us here.[8]

The second notion is that of a *classical valuation*. This is an assignment of truth (T) and falsity (F) to sentences, which disregards the possibility that a sentence might suffer from a failure of presupposition. Given our remarks about *or* and *not*, we know three things about classical valuations:

8. If v is a classical valuation for the language L and A, B are sentences of L, then
   (a) $v(A) = \mathrm{F}$ or $v(A) = \mathrm{T}$,
   (b) $v(\sim A) = \mathrm{F}$ if and only if $v(A) = \mathrm{T}$,
   (c) $v(A \vee B) = \mathrm{F}$ if and only if $v(A) = v(B) = \mathrm{T}$.

Classical propositional logic is clearly *sound* with respect to classical valuations, since classical valuations correspond to rows in the ordinary truth tables. But of course, there may be sentences of L that are assigned T by all classical valuations (*classically valid* sentences of L) which are not theorems of the propositional calculus. (From here on we shall say that a valuation *satisfies* A if it assigns T to A, and that it *satisfies* a set X if it satisfies every member of X.)

[8] In this paper, N will be taken to be a relation of sentences to sentences; the general case of such a relation from sets of sentences to sentences will not concern us here.

In general, no classical valuation will be entirely correct with respect to an actual situation. With respect to an actual situation, the sentences of L are divided into three classes: those which are true, those which are false, and those which are neither. Since a sentence is true if and only if its negation is false, it follows that which sentences are true determines which sentences are false, and hence determines also which sentences are neither true nor false. In other words, in so far as the situation can be described in L, it is determined uniquely by the set of sentences of L which are true with respect to it.[4]

Suppose that G is the set of sentences true in a given situation. What do we know about G? First, no sentence in G has a presupposition that fails. Therefore, the sentences of G can all be true together from the point of view of the classical valuations, and what follows from them from this point of view really does follow. (There are sentences about which the classical valuations are radically wrong, since they cannot accommodate a lack of truth value; but such sentences do not belong to G.) On the other hand, the classical valuations disregard N; hence there are some consequences of true sentences that they overlook. If A belongs to G, and $N(A,B)$, then B is also true; hence also belongs to G. We may sum this up as follows:

9. (a) There is a classical valuation that satisfies G,
   (b) if every classical valuation that satisfies G also satisfies B, then B is in G,
   (c) if $N(A,B)$ and A belongs to G, then so does B.

A set of sentences G for which 9(a) holds we call *classically satisfiable*, and if 9(b) and 9(c) hold for it, we call G a *(necessitation-)saturated* set. So our conclusion is that the set of sentences true in an actual situation is a classically satisfiable, saturated set.

It is at this point that we can discuss the admissible valuations. An admissible valuation is to correspond to a possible situation, such that it assigns T to the sentences true in that situation, F to the sentences false in that situation, and does not assign a truth value to those sentences which are neither true nor false in that situation. On the basis of our present discussion, we can therefore say the following about admissible valuations:

10. An assignment s is an *admissible valuation* for L only if there is a classically satisfiable, saturated set G such that:
    (a) if A is in G, then $s(A) = \mathrm{T}$,
    (b) if the negation of A is in G, then $s(A) = \mathrm{F}$,
    (c) otherwise, $s(A)$ is not defined.

[4] This can be made precise by using the notion of *model,* but the intuitive notions suffice for our present purpose.

We may note that $(\sim\sim A)$ is in G if and only if A is in G; so a sentence has (or lacks) a truth value if and only if its negation does. Also, since G is saturated, if A necessitates B and B is not in G, neither is A. From this it follows by our characterization of presupposition that if A has a presupposition which is not true, then A is neither true nor false.

There is another important point about admissible valuations, which establishes that whatever logic is sound with respect to the classical valuations is also sound with respect to the admissible valuations. This is the point that an admissible valuation represents what is common to a certain set of classical valuations. The assignment s characterized by 10 is identical with the supervaluation induced by the set G in question—which notion is defined by:

11. The *supervaluation induced* by G is the function that
    (a) assigns T to A if all classical valuations that satisfy G assign T to A,
    (b) assigns F to A if all classical valuations that satisfy G assign F to A, and
    (c) is not defined for A otherwise.

And clearly what is common to classical valuations cannot transgress laws that hold for all classical valuations (in particular, the laws of classical propositional logic).

Specifically, the *law of excluded middle* continues to hold:

12. Any sentence of the form $(A \vee \sim A)$ is valid (assigned T by all admissible valuations).

But the *law of bivalence* does not hold: some sentences are neither true nor false.

Reasoning concerning presupposition, implication, or necessitation can be tested through our formal paradigm. With a view to this critical function of the formal paradigm, we shall point out certain features of the notion of presupposition which can easily be demonstrated with its help.

First, the reader will have no difficulty in verifying that there can be distinct cases of necessitation, implication, and presupposition, so that these are three distinct semantic relations. Secondly, we may note that if A presupposes A, then A is never false (it may be true, or neither true nor false). Similarly, if A presupposes $(\sim A)$, then A is never true. If both are the case, if A presupposes a contradiction, then A is always neither true nor false. And if A is presupposed by a valid sentence, then A is valid. (It is clear that the relation N might be such that the language had no admissible valuation: a pathological case which could model only discourse with inconsistent presuppositions.)

III

The first subject to which we shall apply the distinctions drawn above, is *truth*. This is not caprice; the points here made will then play a role in our discussion of the paradoxes of self-reference. The main question before us is: what is the relation between the sentence "It is true that P" and P itself? (We symbolize the former as $T(P)$.) The answer purports to be given by Tarski's principle:

13. $T(P)$ if and only if P.

But what relationship is this "if and only if" meant to indicate? One obvious answer is that it signifies *co-implication*. But this is not necessarily so; for example, one *might* say "If the King of France is bald, then he exists" to signify that the antecedent necessitates the consequent. (This would be a confusing use of 'if . . . then', to be sure, and would most likely indicate a failure to distinguish between implication and necessitation.)

Do P and $T(P)$ imply each other? Principle 13 is used in the "derivation" of the semantic paradoxes, and has also been used to "derive" the law of bivalence. I wish to consider this latter "derivation" here, and shall present it in its shortest form. Given that 13 must be understood to assert a co-implication, we have in particular

14. (a) P implies $T(P)$.
 (b) $\sim P$ implies $T(\sim P)$.

Suppose now that $T(P)$ is not the case; that its denial is the case.[5] Then, by (a) and modus tollens, the denial of P is the case. But this conclusion and (b) lead by modus ponens to: $T(\sim P)$ is the case. So if $T(P)$ is not the case then $T(\sim P)$ is the case. Since our metalinguistic 'if . . . then' is the material conditional, this means that either P is true or $\sim P$ is true: that is, either P is true or P is false. But if the language is a presuppositional language, this conclusion does not hold.

The first, and obvious, reaction to this argument is that the distinction between use and mention has not been observed. Let us try to observe it. Let L be a presuppositional language, and let us form its metalanguage M as follows:

15. (a) sentences of L are first-level sentences of M,[6]
 (b) if P is a sentence of L, then 'P' is a name in M,

[5] This assumption of bivalence for $T(P)$ will be discussed shortly.

(c) 'T' is a predicate of M which is applicable only to names formed by applying quotation marks to sentences of L,
(d) if A and B are sentences of M, so are $\sim A$ and $(A \vee B)$,
(e) A sentence that has any well-formed part beginning with 'T' is a second-level sentence of M,

where '$\vee$' and '$\sim$' are also the disjunction and negation signs of L. We continue to let '$T(P)$' stand for 'T' followed by P in quotation marks. We shall use 'P' and 'Q' to stand for first-level sentences only.

We do not at this point have the full semantics for M, but at least the semantics of the first level must coincide with the semantics of L. But our argument above is easily restated making use of M (and taking as premise only the minimal assertion that P materially implies $T(P)$). But, however we complete the semantics of M, the argument cannot establish that the principle of bivalence holds for first-level sentences, because that is not so.

We have three possibilities before us. We can reject the premises, or we can reject the rules whereby the conclusion is derived, or we can accept the conclusion as formulated by the second-level sentence

16. $$T(P) \vee T(\sim P)$$

of M, but interpret this sentence differently. Let us consider this last alternative first. We remember that a sentence of the form $P \vee \sim P$, which we had been inclined to interpret as saying that P is either true or false, need not be so understood. We could similarly use supervaluations or a many-valued matrix to reinterpret second-level disjunction in such a way that 16 does not say that P is either true or false. The most obvious way to do this is to say that when P is neither true nor false, then $T(P)$ does not have a truth-value either —as opposed to: then $T(P)$ is false. But we shall then have no way of formulating the assertion that a sentence is *not true*. Nor could we add a third level to M in which to formulate this assertion without running into the same problem. That is, assuming that the transition to each higher level obeys the same principle, we would simply get: if P is neither true nor false, then neither is $T(P)$, and neither is $T(T(P))$, and so on. Thus we have here a solution, but it means that M is in some important ways not a model of the metalanguage we have actually been using.[7]

The second possibility, of rejecting the validity of the argument, is not a very pleasant one either. The core of the argument, which can be formulated in M, is

[6] Strictly speaking, the first level of M should be isomorphic to, rather than identical with L; but this makes no essential difference here.

[7] This also answers the question of why we assumed bivalence in the formulation of the argument (see fn. 5).

17. $P \supset T(P)$
$\sim P \supset T(\sim P)$ contraposition
$\sim T(P) \supset \sim P$
$\sim T(P) \supset T(\sim P)$ transitivity
$\sim\sim T(P) \vee T(\sim P)$ def. material implication
$T(P) \vee T(\sim P)$ double negation

and is entirely validated by classical propositional logic. Again, I would not be satisfied if this were rejected; such a rejection might throw radical doubt on our own metalogical reasoning. But, instead of rejecting 17, we might reject the original argument on the basis that from "A implies B" we ought not to conclude that A materially implies B, and that the moves in 17 do not apply to implication proper. I would not care to prevent anyone from introducing a new arrow (this has been a popular and, I would claim, harmless pastime), but I do not really think it would be to the point here.

Rather than turn to such radical departures from the standard logical framework, we may consider the possibility that the argument is not sound. As we originally formulated it, the argument proceeded mainly by modus tollens and modus ponens. Only the latter is valid for necessitation in general. Therefore we will have circumvented the problem very simply by interpreting Tarski's principle not as a co-implication but as a *co-necessitation*.

We can now also complete the semantics of M in a very obvious way:

18. (a) If the first-level sentence P is true, then $T(P)$ is true, and otherwise $T(P)$ is false;
(b) If the second-level sentence A is false, then $\sim A$ is true, and if the second-level sentence A is true, then $\sim A$ is false;
(c) If $(A \vee B)$ is a second-level sentence, then if either A is true or B is true, then $(A \vee B)$ is true, and otherwise $(A \vee B)$ is false.

Clearly, if A is neither true nor false, then it is a first-level sentence, and so is $\sim A$; if B lacks a truth value as well, then $(A \vee B)$ is a first-level sentence. The principle of bivalence holds for second-level sentences. Classical propositional logic remains sound for the language M.[8]

That this semantics satisfies the principles we have adopted is seen as follows: when P is true, so is $T(P)$, and conversely; hence the two necessitate each other. When P is neither true nor false, then $T(P)$ and $T(\sim P)$ are both false, and $T(P) \vee T(\sim P)$ is also false. When P is neither true nor false, then so is $\sim P$, and $T(P)$ is false;

[8] This is easily seen by embedding the valuations into a three-valued matrix in which the negation of the middle value gets T and the disjunction of two middle values gets F.

hence $(\sim P \vee T(P))$ is false: P does not materially imply $T(P)$. There is, finally, an interesting case of presupposition in M: P does not presuppose $T(P)$, since $\sim P$ does not necessitate $T(P)$. But both P and $\sim P$ necessitate $T(P) \vee T(\sim P)$. Therefore P presupposes its own bivalence; we may call this the ultimate presupposition. Moreover, this is rather a stable result, for to show it we need not appeal to the bivalence of second-level sentences nor to the truth definition for disjunction beyond the unproblematic feature that, when one of the disjuncts is true, so is the disjunction. Not all the results of this section have such stability, as we shall see when we leave the relative security of M.

IV

So far we have assumed that there is a neat division between sentences about sentences and the sentences they are about. I do not mean simply that we have observed the distinction between use and mention. I mean that, in addition, this distinction corresponds to a division of the class of sentences with which we have so far been concerned. That is not the same thing: there is certainly a distinction between loving someone and hating someone; yet it is possible to hate someone you love. So we have not only observed the distinction between use and mention, but assumed that no expression is mentioned in the course of its use.

But this assumption is not a necessary one, and this brings us to the subject of self-reference. The famous Liar paradox is the obvious point of departure. The Liar says "What I now say is false." Clearly this is an English sentence, perfectly grammatical. Yet it can be construed as mentioning itself: the Liar could equally well have said, "This sentence is false" or "The sentence which I now utter is false." The use of this sentence involves mentioning it; hence the distinction between using sentences and mentioning sentences, while a perfectly good distinction, does not correspond to a neat division among sentences.

The paradoxical element appears when we ask whether what the Liar said was true or false. (In the following argument, we assume principle 13 only in the weakened sense of "P and $T(P)$ necessitate each other.") If what he said was true, then what he said was the case; but what he said was that what he said was false. So if what he said was true, then it was false. Similarly we can demonstrate that if what he said was false, then what he said was true. In other words, both the supposition that it was true and the supposition that it was false, lead to absurdity.

This conclusion is itself absurd only on the assumption that what he said must be either true or false. But we are now quite used to the failure of bivalence; so we simply say: what he said was neither true nor false. The air of paradox is spurious.

Before we begin to feel too smug about this, however, we must face a second paradox, which I shall call the *Strengthened Liar* and which was designed especially for those enlightened philosophers who are not taken in by bivalence. The Strengthened Liar says "What I say is either false or neither true nor false."

If we now ask whether the sentence is true or false or neither, we find that each of these answers is absurd. For example, suppose that what he says is neither true nor false. Then clearly it is either false or neither true nor false. But then what he said was the case. So what he said was true. And now we seem to be properly caught, our sophistication with respect to bivalence notwithstanding.

One move that one might consider, though unhappily, is to say that there is a fourth possibility. This has first of all the unwelcome consequence of facing us with a Strengthened Strengthened Liar (and so on *ad infinitum*). Secondly, it is our desire to conform to classical logic. The principle that has for us replaced bivalence is:

19. $$T(P) \vee T({\sim}P) \cdot \vee [{\sim}T(P) \,\&\, {\sim}T({\sim}P)]$$

which is a second-level sentence and is valid no matter what P is, for it is a tautology, a theorem of propositional logic. We can deny $T(P) \vee T({\sim}P)$, bivalence, for that is not a tautology, but we cannot deny 19.

However, I have not led you through the labyrinthine distinctions among implication, necessitation, and presupposition merely for its own sake. But before I show how these distinctions may be mobilized to help us see our way through the Strengthened Liar paradox, I must say something about English sentences and our symbols. We use ${\sim}P$ to stand for the denial of P, and so when P is an English sentence, say "Tom is tall," we generally read ${\sim}P$ as "Tom is not tall" or "It is not the case that Tom is tall." But if X is, for example, the Liar sentence "What I now say is false," we see that its denial is not expressed by "What I now say is not false." If the Liar were to utter both, he would have to utter them in succession, so the 'now' would refer to two different times. Hence he would not deny his first statement by making the second, the second referring only to itself. Yet X does have a denial—the Liar's audience, or he himself, may respond "That is not so." Similarly, in deriving the absurdity we argue, for example, that if what he says is not the case, then what he says is false—using 'what he says is not the case' to express the denial of what he says. The exact English words which may express this denial will depend on *who* denies it *when*. This is a clear sign that our symbolism is not nearly adequate to give a complete picture of self-referential language.

But what we can do is to isolate the features of the Liar sentence that play a role in the paradox. In doing so, we may use X to stand for what the Liar says, and $\sim X$ for its denial, however (when, by whom) the denial may be expressed. And then we see that we can take the essential features of X that are appealed to in the derivation of absurdity, to be its relations to the sentence $T(\sim X)$ which expresses its falsity. In all other aspects, we shall assume our formal paradigm to govern inferences involving X, for this reduces the problem to the familiar case. And we shall have succeeded if we can then show how the derivation of absurdity is blocked.

First, we might take X to co-imply $T(\sim X)$ and also $T(X)$, and $T(\sim X)$ to co-imply $\sim X$. This is the first way in which anyone is likely to take the problem, and it leads to absurdity. *This shows* that this way of taking it is incorrect. (In any case, we have already seen that some of these implications do not hold.) Let us now be cautious and take these relations all to be cases of co-necessitation. Then we find that bivalence is needed to derive absurdity; and this dissolves the paradox in the way we indicated informally above.

We could also put this as follows: X necessitates $T(\sim X)$, but so does $\sim X$. Hence, by our definition, X presupposes $T(\sim X)$. We have seen that $T(\sim X)$ cannot be true; therefore, X has a presupposition that is not true. This is why X is neither true nor false. In this way, the distinction between necessitation and presupposition leads to a solution of the paradox.

Thus the Liar paradox seems to fit very nicely into our conceptual scheme. But the fact is that it has one feature that does not fit at all. For X is itself an assertion of falsity: the Liar says "What I now say *is false*." And in the preceding section, when we constructed the metalanguage M, we followed the principle that, although bivalence does not hold generally, it does hold for assertions of truth and falsity. We are here taking the Liar sentence to be a first-level sentence, but there are no restrictions on the form of first-level sentences. In particular, this sentence is an assertion of falsity, and it does not have a truth value.

This I shall call the *basic lesson* of the Liar paradox: even assertions of truth or falsity do not in general satisfy the law of bivalence.

Turning now to the Strengthened Liar paradox, we find a sentence Y which co-necessitates its own falsity-or-truthvaluelessness. That is, Y and $(T(\sim Y) \vee \cdot \sim T(Y) \,\&\, \sim T(\sim Y))$ necessitate each other. Each of the three possible suppositions—that Y is true, that Y is false, that Y has no truth value—necessitates a contradiction. We have here three valid arguments with a common conclusion; and this conclusion is a self-contradiction. This conclusion is demonstrated if one of these three valid arguments must have true premises. This is tantamount to:

20. $T(Y)$ is true, or $T(\sim Y)$ is true, or $\sim T(Y)$ & $\sim T(\sim Y)$ is true.

This looks like the tautology 19, but it is not the same *unless* we persist in the opinion that assertions of truth are themselves always true or false. And this was the basic lesson of the ordinary Liar paradox: that opinion is mistaken.

To put it most perspicuously, 20 is related to 19 as bivalence to excluded middle. We cannot conform to logic and also deny 19, or excluded middle. But we can deny bivalence, and we can also deny even that sentences that begin with 'T' are bivalent. The Strengthened Liar paradox is averted if we hold that $T(Y)$ and $T(\sim Y)$ are themselves neither true nor false. From this it follows immediately that the sentence $\sim T(Y)$ & $\sim T(\sim Y)$ also is neither true nor false.

And we can also give a good reason for holding this. As for every sentence, $T(Y)$ necessitates Y. But, by the tautology 19, $\sim T(Y)$ necessitates $T(\sim Y) \vee \cdot \sim T(Y)$ & $\sim T(\sim Y)$. The latter in turn necessitates Y. Hence $T(Y)$ and $\sim T(Y)$ both necessitate Y; in our terminology, $T(Y)$ presupposes Y. As we have seen, Y cannot be true. This is why $T(Y)$ is neither true nor false. (Similarly for $T(\sim Y)$.)

At this point it may be instructive to see how we would extend our formal paradigm to accommodate the deviant sentences X and Y. We must be careful again not to extend it in such a way that classical logic is violated, for then our own reasoning might be drawn in doubt. First we place X and Y among the first-level sentences. They bear no unusual relations to other first-level sentences; hence we need not extend the relation N. But we must now add to the semantics of M a relation N* of nonclassical necessitation, and say that X and $T(\sim X)$ bear N* to each other. Similarly Y and $T(\sim Y) \vee \cdot \sim T(Y)$ & $\sim T(\sim Y)$ bear N* to each other. Let us call the thus extended language "M*." The admissible valuations of M become the classical valuations of M*. We define 'saturated set of sentences of M*' as before, except that we use N* instead of N. The admissible valuations for M* are supervaluations generated by saturated sets of sentences, as before. Only the sentences belonging to the entourage of X and Y will be affected by this; the others keep their normal truth value (if any). That classical logic continues to hold can be demonstrated as before.

The notion of "level" has been made much less sharp; the relation N* imposes trans-level semantic relations among the sentences. But there is still a clear and distinct *syntactic* notion of level. We can also add a third level, in which we can, for example, express the fact that $T(Y)$ is neither true nor false. We would do this by extending M* in just the way that we extended L to produce M. But let us not deceive ourselves: we shall not get to a point where we can say everything relevant, and yet not have any presuppositions that could fail. To be presuppositionless may be a regulative ideal in philosophy, but it is not an achievable end.[9]

In conclusion, I should like to describe briefly a further kind of paradox, and attempt to apply my analysis to it. Epimenides the Cretan is reported to have said that all statements by Cretans are false. Clearly, what he said cannot be true. For what he said was said by a Cretan, and hence he has implicitly asserted its falsity. But we can consistently hold that what he said is false. This just means that *something* said by some Cretan is not false. And this is not as implausible as Epimenides seems to have thought. But, as Church has pointed out, whether or not some statement by a Cretan is not false, is a contingent matter. In particular it entails that the statement of Epimenides which I have just described is not the only one ever made by any Cretan. So the world could have played a neat trick on us: this could have been Epimenides' first and last statement, and all other Cretans could have been entirely dumb. In that case, neither could we have held that what he said was false.

This paradox we shall call the *Weakened Liar*. Epimenides said in effect that all other statements by Cretans are false, and also that his own (this very) statement is false. Let his sentence be Z and let the sentence that expresses that all other statements by Cretans are false be Q. Then Church's point is that the denial of Z necessitates $\sim Q$.[10] But Z necessitates $T(Z)$, which necessitates a contradiction; hence Z also necessitates $\sim Q$. Therefore, $\sim Q$ is a presupposition of Z; and if it fails, Z is neither true nor false. But if $\sim Q$ is true, the case is different: Z and Q are such that $\sim Q$ necessitates $\sim Z$. Therefore if $\sim Q$ is true, then Z is false.

This final example intends to show how our analysis can be applied to members of this family of paradoxes other than the two for which it was developed explicitly. On the other hand, we have not supplied a general theory of self-referential language. This is the familiar lament that we have no (sufficiently general) formal pragmatics. But the paradoxes of self-reference have been a major ob-

[9] That we envisage only finitely many levels here is not a necessary limitation, and not essential to this point.

[10] In the context of the fact that the statement is made by a Cretan; in the present argument, *necessitation* and *presupposition* are relativized to this assumption.

stacle to such a pragmatics, and we shall be satisfied if we have shown that this obstacle at least is not insuperable.

BAS C. VAN FRAASSEN

Yale University

NOTE. The standard discussions of *presupposition* are those of P. F. Strawson in his *Introduction to Logical Theory* (London: Methuen, 1952), and "On Referring," *Mind*, LIX, 235 (July 1950): 320–344. Strawson's account has been critically discussed a.o. by W. Sellars, "Presupposing," *Philosophical Review*, LXIII, 2 (April 1954): 194–215, and G. Nehrlich, "Presupposition and Entailment," *American Philosophical Quarterly*, II, 1 (Jan. 1965): 33–42. The former was answered by Strawson in the same issue of that journal, pp. 216–231, and the latter corrected by Nehrlich himself in "Presupposition and Classical Logical Relations," *Analysis*, XXVII.3, 117 (January 1967): 104–106. The literature on *implication* is now too voluminous to be summarized; we refer only to A. R. Anderson and N. D. Belnap, Jr., "The Pure Calculus of Entailment," *Journal of Symbolic Logic*, XXVII, 1 (March 1962): 19–52. Correspondence with Peter Woodruff, Wayne State University, has helped me to become clearer on the distinction between implication and presupposition.

The basic idea for our treatment of *presupposition* was suggested by Karel Lambert, University of California at Irvine, to whom I am much indebted, in a discussion of section VII of this author's "Singular Terms, Truth-value Gaps, and Free Logic," this JOURNAL, LXIII, 17 (Sept. 15, 1966): 481–495. The language of free logic as there described is a presuppositional language in the sense of sec. II of this paper. The notion of *supervaluation* in this paper is a generalization of that presented in "Singular Terms, Truth-value Gaps, and Free Logic." Supervaluations have since been used in R. Meyer and K. Lambert, "Universally Free Logic and Quantification Theory" (forthcoming), and in B. Skyrms' comments on J. Pollock's "The Truth about Truth" at the APA (Western Division) meetings in Chicago on May 4, 1967 (see below).

The distinction between excluded middle and bivalence, apparently first made by Aristotle, was introduced in this century by the Polish logicians: see S. McCall, "Excluded Middle, Bivalence, and Fatalism," *Inquiry*, IX, 4 (Winter 1966): 384–386. That we are in good company here is witnessed by Quine, who mentions Paul Weiss, Yale University, as having been brought "to the desperate extremity of entertaining Aristotle's fantasy that 'It is true that *p* or *q*' is an insufficient condition for 'It is true that *p* or it is true that *q*'." (*The Ways of Paradox*, New York: Random House, 1966, p. 21). The nature of validity and the consequence relation in presuppositional languages is explored in my "Presuppositions, Supervaluations, and Free Logic" in the forthcoming Festschrift in honor of Henry Leonard (ed. by Karel Lambert).

Section III is an improvement and extension of sec. VIII of "Singular Terms, Truth-value Gaps, and Free Logic." The three-valued matrix mentioned in sec. III was suggested by footnote 23 of E. Sosa, "Presupposition, the Aristotelian Square, and the Theory of Descriptions" (mimeographed, University of Western Ontario, 1966); the question of the adequacy of such matrices for classical propositional logic is discussed in A. Church, "Non-Normal Truth-Tables for the Propositional Calculus," *Boletin de la Sociedad Matematica Mexicana*, X, 1–2 (1953). I should like to thank John Heintz, University of North Carolina, for helpful suggestions concerning the use of this matrix.

The literature on the paradoxes of self-reference is also huge; this paper has profited most from A. N. Prior, "On a Family of Paradoxes," *Notre Dame Journal of Formal Logic*, II, 1 (January 1961): 16–32. Also suggestive were N. Rescher, "A Note on Self-referential Statements," *ibid.*, V, 3 (July 1964): 218–220; and D. Odegard, "On Weakening Excluded Middle," *Dialogue*, V, 1 (September 1966): 232–236. Bryan Skyrms, University of Illinois at Chicago Circle, has developed independently a solution to the Strengthened Liar paradox through the use of supervaluations, which he presented in the paper mentioned above. His account is in some ways very similar to ours, but the syntax of the artificial language used is quite different; it may lead to a much finer analysis of the self-referential language used in the paradox than we have provided.

Finally, I should like to acknowledge my debt to stimulating discussions and correspondence on the semantic paradoxes with V. Aldrich, University of North Carolina, and G. Nakhnikian, Wayne State University, and to thank R. Clark, Duke University, and P. F. Strawson, University College, for their encouraging comments on an earlier version of this paper.

BOOK REVIEWS

Elements of Mathematical Logic. JAN ŁUKASIEWICZ. Translated by OLGIERD WOJTASIEWICZ. Oxford: Pergamon Press, and New York: Macmillan, 1964. xi, 124 p. $6.50.

Łukasiewicz's *Elementy logiki matematycznej* first appeared in 1929 in the form of mimeographed lecture notes. In 1958 a printed edition appeared, edited by J. Słupecki. The present volume is a translation of this printed edition.

English-but-not-Polish-speaking readers owe the translator a very great debt for bringing this remarkable little work within their comprehension. After an introduction which is partly historical, partly apologetic, partly explanatory (it brings out the differences between rules and theorems), the sentential calculus is systematically developed, in the book's longest chapter, from the author's well-known set of three axioms *CCpqCCqrCpr*, *CCNppp* and *CpCNpq*. The consistency, independence, and completeness of this basis is then demonstrated. In chapter IV quantifiers are introduced, but only as binding sentential variables; and finally there is a sketch of the author's well-known formalized version of Aristotelian syllogistic.

By contemporary standards, this does not take us very far; but what is done, is done with an extraordinary elegance, clarity, and perfection. It is a pity the book is not cheaper, for it is still as good an introduction as can be found to the field which it covers, and deserves to be widely read and used by students.

It is perhaps also a pity that the translation is from the 1958 rather than the 1929 version, as Słupecki's alterations haven't always been improvements, and in any case have been a little unfortunately carried out. He has omitted a supplement "On Reasoning in the Natural Sciences" which English readers really ought to have, as it gives a brief lucid presentation of the hypothetico-deductive view which in English-speaking countries we tend to associate not with Łukasiewicz but with Popper. Słupecki has also, in chapter III, replaced