merits: that of groundedness and that of contingency of the truth conditions of an utterance.

2.3. The theory

In order to outline the main features of his theory, Kripke considers an arbitrary language L, which is not, of course, a real natural language such as Italian: to address the latter directly would be a mammoth task, which, in addition to clashing with an infinity of other problems that have nothing to do with the notion of truth, would not allow us to focus on the main features of the proposal. More simply, Kripke asks us to imagine that L is a language that shares with natural languages the characteristic of being rich enough to express its own syntax and semantics, which is the characteristic responsible for all the problems we have seen.

For this purpose, we may assume that L is essentially what we obtain by adding a unary predicate *T,* which we shall read as 'is true', to a common predicate logic language such as the one we defined at the beginning of section 1.5, which we shall denote by Lt. For obvious reasons of simplicity, we shall assume that Lø is devoid of any modal operator. However, we will assume that Lp has all the resources to satisfy the requirement of syntactic expressiveness mentioned above, and in particular that it is possible to associate any well-formed formula of Lo with a name belonging to the vocabulary of Lt itself. This requirement may be met in various ways. In Italian, perhaps the most common way to assign a name to any utterance is to write that utterance in inverted commas. Another way, followed extensively in these pages, is to assign each utterance a certain numeral that uniquely identifies it, rather like the digits printed on a passport uniquely identify each citizen. So, for example, we can speak of the utterance 'The number of planets is greater than 7' or of the utterance (72): in both cases we are not *using* the utterance in question; we are *naming it,* and in this way we can attribute properties to it by using certain predicates (e.g. the predicate *'and* true'), just as we can attribute a property to a person by calling him by his name and applying a certain predicate to that name (e.g. 'he is a philosopher'). Returning to Lø, Kripke assumes that formula names are fixed in the second way, which is somewhat simpler as it does not require the addition of functors that, like inverted commas, are not governed by the laws of the logic of predicates. In practice, this means assuming that the vocabulary of Lp includes the entire vocabulary of the common language of arithmetic, and that a procedure has been defined for assigning each well-formed formula of Lø a specific number, e.g. an 'arithmeticisation' procedure of syntax similar to that introduced by Kurt Gödel for the proof of his famous Incompleteness Theorem of 1931.

Let Lo therefore be a language of the type just described, and let L be the extension of Lø obtained by adding the predicate *T* to the latter's vocabulary. The arithmetisation procedure can obviously also be extended to L. If *A* is a well-formed formula of L, we shall write " *A° "* to indicate the numeral of its "gödelian", i.e. the number assigned to *A* by the procedure in question. In this way, we can intuitively read *T "A"* as a formula that says, in L, that the formula *A* of L is true (and *T"-A"* as the formula that says that the negation of *A is* true*,* i.e. that *A* is false). On this basis, the first step towards the formulation of the theory will evidently consist in the definition of the models of L. After that, the crucial step will be the development of precise criteria for identifying the models of L that provide an *adequate* interpretation of *T,* i.e. an interpretation that justifies the intuitive reading that we wish to attribute to *T("is* true").

With regard to the first point, we can safely rely on11the standard definition given at the beginning of section 1.5, with one variant: that the interpretation assigned by a model œp to the predicate *T* may be a *partial* function*,* i.e. defined only for certain elements of the domain D. In other words, can(Ę will assign to certain elements of D one of the two values V or F, as for any unary predicate, but it is possible that there are elements to which ctp(Ę does not assign any value. In the extreme case, it is also possible that the function on(Ę is undefined for *all* elements of D. Otherwise, however, we can assume that the models of L behave exactly as in section 1.5, which is the same as saying that their restrictions to Lø will be entirely standard models.

Intuitively, the reason for the variant just described should be obvious. Insofar as *T* must represent the predicate 'is true', requiring ctp(Ę to be a total function would be tantamount to assuming that *T* obeys a form of the bivalence principle. But we have already seen that in the case of certain statements, such as that of the liar, bivalence is sufficient to generate a paradox. Therefore, if we wish to hope that the theory will return us a truth predicate that behaves coherently, we must leave open the possibility that the application of *T* to the name of a formula *A* resolves in a neither true nor false statement, and this means that the application of on(to *ao("A")* must be able to be undefined. The important thing to note at this point is that by relaxing the notion of a model in this way we can no longer rely on the standard characterisation of clauses for the evaluation of well-formed formulae that we summarised in (46)-(50). Those clauses were formulated under the assumption of bivalence, so that the untruth of a formula with respect to a given model always coincided with its falsity. If, however, certain pre-

In this case, the predicate F corresponds to partial functions, the clause concerning atomic formulae must be more explicitly reformulated so as to distinguish between cases in which a formula is not true with respect to the given model and cases in which it is false.

atomic formula is neither true nor false, the clauses relating to statements composed by means of logical operators will in turn have to be reformulated so as to provide precise instructions on how to evaluate those formulae that contain parts, i.e. sub-formulae, that are neither true nor false. In short: the possibility of *T* being interpreted as a non-bivalent predicate immediately results in the need to define a non-bivalent semantics for the language L.

We have already observed, with reference to the semantics of quantified modal logic, that this task presents difficulties. However, there is no shortage of proposals, and Kripke himself considers that in this context the choice of one solution over another is largely arbitrary. The only important condition is that the chosen semantics satisfies an intuitive requirement: the more determined the interpretation of the predicate F is, the more determined the truth conditions for the language L are. More precisely, the models of the semantics must satisfy the following 'stability requirement':

(74) If op differs from bp merely because up(is defined for some elements of D for which bp(is not defined, then what is true in [3p remains true in top and what is false in (ip remains false in op.

For the sake of completeness, here are the clauses defining one of the non-bivalent semantics considered by Kripke, due to logician Stephen C. Kleene. These clauses are a natural extension of the standard clauses given in (46)-(50) and differ from them only in the fact that, for each well-formed formula J, one gives both the conditions under which it is true with respect to a given model cp (op , and the conditions under which it is false (with J. It is understood that in cases where both conditions are not met, the formula itself is to be considered neither true nor false.

(75) ctp 1= *Pti . . .t* if and only if *ao(P)[iio(ti),* ..., crp(/,)) = V
ao Pti ... . *tao(Pyao(ti)*, ..., ctp(i,)) = F
(76) 'zp i= -*A* *ao A*
ctp -*Aao A*
(77) op 1= *A ri B* *no A* and cap *B*
op *A* ri *B* op =J *A* o op =J *B*

(78) up t= *VxA* [ip 1= *A'* for each /-variant §p of ctp
$\alpha_D \Dashv \forall xA$ [ip =i *A',* for some i-variant §p of op [ip
(79) $\alpha_D \models \exists xA$ 1= *A* ',. for some /-variant §p of vip [lp =J
$\alpha_D \Dashv \exists xA$ *A'* for any f-variant §i, of maj

Let us now come to the second step in the formulation of the theory, the crucial and profoundly innovative one: the identification of the models of L that form an *adequate* interpretation of F. Which are, among the infinite models through which we can interpret L, those in which *T* pro- prly *represents* the predicate "is true"? Evidently we would like to answer that the models in question are no more and no less than those models in which *T* reflects the truth conditions for the statements of L, i.e. those models ctn which satisfy the following "adequacy condition" for every well-formed formula *A:*

CA *an A if* and only if nn T *"A"*
"p < *A* if and only if "p < T "*A"*

The problem is that there is no guarantee that such models exist. If there were none, we would have confirmation of the Tarskian thesis that a 'semantically closed' language such as L is inherently incoherent, even in the absence of the bivalence principle. If, however, we succeed in proving the existence of at least one model that satisfies CA, then we have in this way demonstrated that it is possible to have a real theory of truth even for such languages. Well, Kripke's great contribution lies precisely in the demonstration of this fact. More precisely, it lies in having provided a demonstration of this fact in the light of which it is possible to give a convincing explanation of the problematic behaviour exhibited by the notion of truth in natural languages of which L is a formal image. (On a purely mathematical level, the existence of a model conforming to CA for a semantics quite similar to that defined in (75)-(79) had also just been demonstrated by Richard L. Martin and Peter W. Woodruff in an essay entitled *On Re- presenting "True-in-L" in L,* already available at the time of Kripke's article, although published only the following year, However the existence of

this model did not allow much to be said except, precisely, that Tarskian pessimism was not justified on a mathematical level).

Before describing the maternity structure of the demonstration, which in some respects consists of a veritable "construction" of models for L-forms in AC, it may be useful to anticipate the intuitive idea with reference to a natural language such as Italian. Kripke himself invites us to consider an imaginary situation that contains all the essential elements. Suppose - says Kripke - we had to teach the meaning of the predicate 'is true' to a person who does not possess it. How would we proceed? Obviously we would tell our interlocutor that a competent speaker of Italian is authorised to apply the predicate 'is true' no more and no less than to those utterances that the speaker himself is willing to assert. This answer reflects, with respect to the Italian language, exactly the same intuition that finds expression in the condition of CA adcguacy for the language L. And we can immediately guess what its effects are. First of all, since our interlocutor is willing to assert, say,

(ß0) The snow is white.

our answer will immediately put him in a position to assert:

(ß1) "The snow is white" is true.

We can instead assume that at this point he is able to apply (or not apply) the

competitions) correctly the predicate 'is true' to all those utterances of which it was in
able to comprehend the meaning before our conversation began, i.e. those utterances that according to Tarski should belong to level 0 of the linguistic 'hierarchy': those that concern the ex/ralingui- stic reality. This is the immediate effect, so to speak, of our answer. But what about the utterances belonging to the other levels'? If he has fully understood our instructions, now that he is willing to assert (81) our interlocutor should also be willing to apply the predicate "is true" to that same assertion, and therefore to assert

(82)' 'Snow is white' is true' is true.

Indeed, we may assume that this consideration applies to all assertions containing the predicate 'is true' that he is willing to assert on the basis of the previous consideration. That is, our interlocutor should know how to behave not only in front of all level 0 utterances, but also in front of all utterances belonging to level 1 of the Tarskian hierarchy.

But if this is the case, then we can repeat the reasoning also with reference to (82) and all utterances of the same level: Our interlocutor should know how to apply "is true" also to utterances of level 2. It is understood that at this point the procedure can continue indefinitely so as to allow the evaluation of any utterance belonging to any level n + 1 of the hierarchy, and this is the second important consequence of our initial answer. In addition, it is reasonable to suppose that at this point our interlocutor will also be willing to assert utterances containing the predicate "is true" without explicitly referring to utterances belonging to a particular level, such as

(83) Some statements are true.

that our interlocutor will be able to infer from (S1) or (S2) by csistcnzialc generation. Now, there is no reason to think that our interlocutor will be able to make a decision at this rate with regard to the utterances of the Italian language containing the predicate "is true", which is why it would be unreasonable to force him to accept the principle of bivalence. It cannot be ruled out that within a short time he will be confronted with utterances that, like the liar's predicate, are not so *predicated* on the truth of the level 0 utterances from which he started, and that he will therefore not know what to do. Nevertheless, we can assume that in most cases his use of the predicate 'it is true' is perfectly in accordance with that of a competent speaker.
Anz-i and this is the final effect of our answer- there is to think that a/

/imite of this process his use of the predicate coincides exactly with that of a competent speaker.

Kripke's demonstration is in essence a forinal reconstruction of this type of reasoning with respect to language L, with a ma- tematic confirmation of the final consideration. Let us begin by considering a model of the language L that interprets *T* as a completely empty function, extend it to a model that interprets *T as* limited to all (and only) the Gödelians of the utterances of Lø, and progressively extend this model by saturating the interpretation of *T* on the basis of the previous interpretation. The result will be a model of L that interprets P as a function capable of attributing the value V or the value F to all statements that are based on the initial model of the series for their truth conditions. And a careful examination of the mathematical properties of the series will reveal that this model has the characteristics we are looking for: the interpretation of *T* that emerges is fully in accordance with the adequacy condition CA.

Let us therefore see in more detail how one can construct a series of the type just described'. Let D be a prefixed doininium of objects that includes all natural numbers, and let etc be a certain model of Lø on D that satisfies all the requirements of a standard model of arithmetic. In particular, ep will assign to each numeral0 of Lt the corresponding number in D and to each arithmetic predicate of Lø the corresponding function on D (possibly extended in an arbi- trary way to elements of D that are not numbers). Since D is fixed, we also agree to omit the index 'D' and simply speak of the model

n. We now define, for each ordered pair (Dp, Dt) of subsets of D that do not have elements in common, a corresponding model of L on D, which we denote by ct[D2, Dr]. This model coincides exactly with ct as far as the symbols of Lt are concerned, and interprets the additional predicate P in accordance with the following general conditions, i.e. valid for
every clcmcnto d of
D:

(84) o[D' Dt](Ȩ(d) = V if and only if d C Dt
"[Dç, Dy](Ç(d) = F if and only if d C Dp

In other words, ct[Dç, D¢] is the model of L that is obtained from n by treating Dy as the este0sion of *T,* i.e. the set of those elements in D of which *T* is true, and D¢ as the *countertension* of *T,* i.e. the set of those elements of which P is false. Let us now define a function that to each model ct[Dç, D¢] associates a moiety ct[Dp', Dr'] as scguc:

(85) Dt' = the set of elements of D that are gödelians of the true formulae of L with respect to 'i[Dç, Dp];
D,' = 1 set of elements of D which are gödelians of the formulae of L Palse with respect to ri[Dç, Dp] or which are not the gödelians of any form of L.

In other words, Ø associates with ct[D2, Dt] that model which interprets the predicate *T* by means of a function that faithfully reflects, in the (partial) attribution of the values V and F, the (partial) conditions of truth and falsity determined by o[Dv, Dt] itself. (The decision to put in Dt' the elements of D that are not among the Gödelians of the formulae of L is entirely arbitrary, but reflects the idea that P only applies correctly to well-formed formula names; even in Italian there is not much sense in saying that Saul Kripke is true, or that the number of planets is true, and things will be much simpler if instead of leaving the question open we decide once and for all to treat such cases as false). At this point it is not difficult to realise the significance of this co-struction with respect to the adequacy condition CA: in order for a model ct[Dç, D¢] to satisfy this condition, it will in fact be necessary and sufficient that we have œ[Dç', Dp'] = ct[Dy, Dç]. In other words, it will be necessary and sufficient that œ[Dp, Dt] be a *fixed point,* in the mathematical sense of the term, of the function $. And that it *must* have fixed points is precisely what can be shown by applying the intuitive reasoning illustrated above, i.e. by considering the *limit of* a series of models in which the extension and the counter-extension of the predicate P progressively increase.

At the financial level - and here we are forced to assume a certain dimesti-chczza with the theory of transfinit-i ordinals the series in question may be

The construction is based on notions of set theory and is included here for the sake of completeness. Readers unfamiliar with these notions may proceed directly to the next section.

identified with the set of all models ct; of L emerging from the following inductive definition, where g is any ordinal number:

(ß6) r = 'i[Ø, Ø] if d = 0
(87) " = 'i[Dç', Db'] if d = d+1 and ";= 'i[Dç, Dt]
(88) n = o[Ut< D,t', U,.,Dy'] if ë is a limit and ct= o[Dtt, Dy] for each d < E

Intuitively, the first model of the series, str', corresponds to the conditions in which our interlocutor was before he addressed us: his use of the predicate 'is true' was null, i.e. he determined an extension and a counter-extension of this predicate equal in each case to the empty însîeme Ø. The model o i corresponds to the conditions of our interlocutor as a result of our explanation: the extension of his use of 'is true' contained at that point all the level 0 utterances he was willing to assert, and the counter-extension all the level 0 utterances he was willing to deny. Similarly, o corresponds to the conditions in which our intrlocutorc found himself at the moment when, as a result of a new application of our explanation, the extension and counter-extension of 'is true' had also extended to all level 1 statements. In general, o.. i corresponds to the situation that arose after or applications of our explanation. After that we can think that œ" corresponds to the conditions of our interlocutor once he has realised how this procedure can a/ ***limit*** be iterated an infinite number of times (ui is the first infinite ordinal): the extension of "is true" will include every utterance declared true in at least one of the previous models, and the counter-extension every utterance declared false in at least one of the models preceding. The definitions in (86) 88) merely make this explicit by referring to1 predicate *T* of L, and not only up to1 achievement of ci" rna for the whole series of transfinite ordinals.

The existence of an n, such that $(o,) = e, is at this point a simple consequence of the fact that the set of well-formed formulae of L has a certain cardinality: by dint of extending the extension and the counter-extension of *T,* at a certain point in the process we will have *exhausted* the formulae classified as true or as false, and that point will be by definition a fixed point of $. More precisely, the existence of a fixed point for the series defined in (86)-(88) follows from the cardinality of L together with a1 the fact that, since the semantics in (75 79) satisfies the stability requirement (74), the series thus defined *is conservative,* i.e. it obeys the following *monotonicity* principle:

(89) If 'i[D*y, D*¢] is an extension of 'i[Dy, D¢], i.e. if D*ç includes every element of Dy and D*t includes every element of Dt, then, for each bcn-formula *A* of L, n[D" Dt] *A* only sc o[D*v, D*t] *A* if n[D*" D*r] =i *A.*

2.4. Applications and limitations of the theory

Let us therefore recapitulate. Kripke's theory consists of two main parts. The first is the definition of a non-bivalent formal semantics for a language, L, which shares with natural languages the characteristic of being sufficiently rich to express its own syntax. This part of the theory has no innovative features, to the extent that the truth conditions for the formulae of L are adopted in a fundamentally arbitrary way from a semantic theory - that of Kleene - which had wide application in the 1970s (for example, for the treatment of phenomena such as vagueness or the lack of reference, including the problems mentioned in relation to semantics for lc modal quantificatc logics based on variable quantification domains). In fact, Kripke also discusses different options, including the 'supervaluational' semantics due to Bas van Fraassen, but the question of *which* semantics is best for such a language remains in the background, as long as the stability requirement applies. The second part of the theory is the one that justifies the title of Kripke's article, because it is there that it is a matter of showing that among the many motifs of L there are some that authorise a reading of the predicate *T* as a true predicate of *truth,* and thus the analogy between L and a typical "semantically closed" language such as Italian. It is this second part that meets Tarski's challenge. And it is this part that offers a new perspective and for

certain revolutionary aspects from which to approach the study of the concept of truth and the problems that plague it. It remains to be seen how this fits in with the diagnosis of the problems we have sornmarily summarised in section 2.2, and especially to what extent it can be claimed to have solved them.

First a clarification . We have just seen that on a technical level, the main result consists in identifying suitable models with those models that correspond to a fixed point in a series of the type defined in (86)-(88), the existence of which is guaranteed by the fact that it is a conservative series. It is evident, however, that the series we have constructed is only one of many that fulfil this requirement. For example, instead of starting the series with a cut model that assigns to F an entirely empty extension and counter-extension, i.e. by identifying ct with ct[Ø, Ø], we could have started with a model that classified (arbitrarily) some elements of D in one or the other set. To come back to the case of our Italian interlocutor, this possibility corresponds to the idea that his use of "is true" was not *completely* null: for some reason, he already knew that this predicate applies correctly to certain utterances, for instance certain utterances of level 0. Now, it is not difficult to realise that even starting the series in this alternative way, and defining the subsequent steps as in (87) and (88), the result would still be a conservative series, and therefore we would still have reached a fixed point. This means that Kripke's procedure actually allows us to show the existence, not of one, but of a multiplicity of models in which *T* represents the truth. If we wish, once we have reached a fixed point we can start again, arbitrarily adding some elements to the extension and counter-extension of that fixed point and using the model thus obtained as the initial element of a new series that will end in a new fixed point. The question therefore arises: *which* of these models-i which fixed point-should we favour the choice of an appropriate model for L, assuming ct is an appropriate model for Lt?

For Kripke, the answer is relatively uninteresting. It can be demonstrated that what is obtained by starting with e[Ø, Ø] is the *minimal* fixed point, i.e. such that every other fixed point is an extension of it (in the sense defined in (89)). It can also be shown that there exist *maximal fixed points,* i.e. such that their extensions do not lead to further fixed points. And between these two extremes we have a series of more or less "rich" fixed points, including those that Kripke calls

This clarification is addressed to the reader who has delved into the for- mental details of Kripke's the- ory illustrated at the terrine of the previous section.

intrinsic fixed points. fixed points at which no formula receives a different truth value from that which it receives at other fixed points. Kripke finds the minimum fixed point to be the most natural choice, as is also suggested by the imaginary situation of the person who initially does not know the meaning of "is true", and also has some sympathy for the largest intrinsic fixed point, which proves to be unique and has the interesting characteristic of providing the richest interpretation of *T* that does not depend on *arbitrary* decisions. But it is not on these applications that Kripke invites us to reflect (although it is precisely here that lies the crucial difference to Martin and Woodruff's result cited above: in the terminology just introduced, the adequate model they had demonstrated by different methods essentially corresponds to a maximal fixed point, hence to one of the least 'natural' models, as it is safe from arbitrary decisions). Rather, the important fact for Kripke is that this multiplicity of solutions allows him to articulate with precisions some of those conceptual distinctions that, as we saw in section 2.2, were largely absent in the theories developed up to then. We will only consider the two main cases we have discussed, but they should suffice to illustrate the explanatory potential of the map that has emerged.

The first case concerns the variety of those utterances which in one way or another are problematic purely by virtue of their self-referential form, such as the classical liar, (64), who says of himself that he is false, or what we might call the assertor, (67), who says of himself that he is true. As we have noted, there is a big difference between the two cases: the first is true if it is false and is false if it is true, so that it is impossible to assign a definite truth-value to it; the second is true if it is true and is false if it is false, and to assign a truth-value to it would be arbitrary. Of course, the language L does not contain a literal translation of the two statements in this sentence, since its vocabulary does not include indicative expressions analogous to the word 'this' which appears in both (64) and (67). Nevertheless, L contains formulae that possess exactly the same semantic characteristics. For example, let us suppose that *P* is a syntactic predicate whose interpretation (fixed by the standard model of Lø) assigns the value V to a single element of the domain, namely the Gödelian of the following formula:

(90) *4x(Px -+ -Tx)*

Since (90) says that the *P is* not true, and since the only *P* is precisely its Gödelian, it is clear that we are faced with a formula that says of itself that it is not true, like the liar (indeed, like the strengthened version of the

mentor, which does not depend on the identification of 'not true' with 'false'). Similarly, if the interpretation of the syntactic predicate *Q* assigns V only to the Gödelian of

(91) Vx(Qx -+ *Tx)*

this formula says of itself that it is true, just like the assertor. Well, with reference to such cases, Kripke's theory provides a very clear explanation of the relative similarities and differences. The similarity is that neither of these two formulae is semantically *fundamental* in a sense that we can now precisely define:

(92) A well-formed formula *A* is well-founded if and only if *A* is either true or false with respect to1 minimum fixed point

(from which it follows, due to monotonicity, that a well-founded formula has the same truth value at *all* fixed points). On the other hand, it is easy to realise the difference: the liar's formula, (90), is *never* evaluated as true or false at a fixed point; the assertor's formula (91) is true (or false) at *any* fixed point that includes in the extension of *T* (or in its counter-extension) its gödelian. So the liar is paradoxical, the assertor is not. By specifying a little better the way in which L succeeds in expressing its syntax (which we had to gloss over in order not to make the presentation too heavy), other important differences and similarities we mentioned in section 2.2 could also be characterised in this way. For example, it turns out that the formula corresponding to the disjunction of the classical liar and the assertor, which is in- tuitively not false but can be considered true, has precisely the characteristic of being true at an intrinsic fixed point. And it turns out that self-referential but completely harmless statements such as (69) and (70), which simply say how many words they consist of, are even well-founded. It goes without saying that these definitions do not yet define a complete taxonomy, but it is clear that we are on the way to a rather accurate classification (to whose refinement the subsequent literature has devoted ample resources).

These examples also shed light on the second point emphasised by Kripke in his diagnosis of the problems connected with the concept of truth: the *contingent,* in many cases accidental, nature of such problems. With reference to the liar, for example, it is sufficient to assume that in formula (90) *P* is not a purely syntactic predicate in order to realise how the unfoundedness of this formula cannot be determined a priori. If the interpretation of *P* corresponds to that of the Italian predicate "it is an utterance on the blackboard", e.g. "it is a statement on the blackboard", for example, then the interpretation of P is not a purely syntactic predicate.

pio, then (90) says what (71) said in Italian: that the utterances on the blackboard are not true. And the truth or falsity of this utterance in the minimal fixed point, and before that its self-reference, depends on the exact composition of the extension and counter-extension of *P* in that model, and thus in the initial cut model: it depends on which (other) utterances really appear on the blackboard according to the model. Identical is the case of the Cretan of 6pis/o/a *to Titus,* (63), which corresponds to the hypothesis that the interpretation of # reflects that of the predicate 'was asserted by a Cretan'. And similar is also the case of the paradoxical 'circles' illustrated by the pair (65}-(66) or their strengthened version, which in L we could again represent by formulas such as (91) and (90), re- spectively. If *P'*s interpretation only assigns the value V to the Gödelian of (91), and *Q*'s interpretation only assigns the value V to the Gödelian of (90), then both formulae are unfounded and, more precisely, paradoxical: no fixed point will give them a finite truth value. Sc however, at least one of *P* and *Q* receives a different interpretation, things cainbiano and both formulae may turn out to be well-founded.

These are only examples, but enough to illustrate the effectiveness and explanatory power of Kripke's theory. If we go back to the historical context referred to at the beginning, we can see how the publication of *Outline ofa Theory of Truth* was received with great interest not only by those in the industry, but also by those who harboured serious doubts as to the possibility of arriving at a coherent ana1ysis of the concept of truth and its use in the context of languages not domesticated to the rigid Tarskian hierarchy. It is not an exaggeration to say that, from this point of view, 1976 marks a watershed in the logical-philosophical reflexivity on these issues just as 1959 marks a watershed in the study of to- dal logics. Certainly, this is only a *reut/ine,* as the title states, i.e. a soirmary formulation, and it is a pity that Kripke never produced the more complete version he announced at the beginning of the article. Nevertheless, in the space of a short time Kripke's oui/rue has been subjected to very sophisticated applications and developments, on a strictly logical-mathematical as well as philosophical level, and from the very beginning variants and alternatives have multiplied, which beyond the details confirm the revolutionary impact of this work. (Perhaps the most significant example of this is the so-called "revisionist" theory proposed by Hans Herzberger and, independently, by Anil Gupta and Nuel Belnap, authors in 1993 of the powerful *The Revisíori Theory of Truth.* But the debt to Kripke is also evident in the work of authors such as Van McGee, Aladdin Yaqïib and Keith Simmons, and in the more recent theories advanced by Hartry Field, Tim Mau- dlin, Graham Priest and others, as well as in the Assyrian

treatments of the predicate of

truths initiated by Solomon Feferman in the late 1970s and extensively studied in the following two decades).

At this point, we can conclude with a general remark on the scope of the theory. Kripke himself did not hesitate to use cautionary words in this regard, explicitly stating that he did not consider it to be a definitive solution to all problems. However, there is *a* problem whose failure to solve it could be seen as an indication of an inherent limitation not only of the theory as Kripke sketched it, but of the entire approach on which it rests. And it is a problem that concerns the basic question: can we really consider L, with its beautiful semantics full of fixed points, in the same way as a typical 'semantically closed' language such as Italian? Can we really say that we have found, by studying L, a demonstration of how one can coherently speak a language capable of expressing within it all the truths that concern it?

Unfortunately, the answer is not entirely afferent. As Kripke pointed out in the final pages of his article, the coherence of L necessarily requires a sacrifice on the expressive level. To realise this, it is sufficient to consider what happens in the case of an utterance that theory classifies as paradoxical, such as the classical liar. If *A* is a formula of L expressing such an utterance, we know with certainty that *A* cannot be true with respect to an adequate model of L, i.e. with respect to a fixed point œ . We also know that *A* cannot be false, and this is equivalent to saying that with respect to ct; neither can its negation be true (as can be veri- fied by applying clause (76) for the evaluation of negated formulae). Now, in the language L, these two facts can be expressed through the following well-formed formulae:

(93) *-T'A'*
(94) *-T"-A"*

However, if ct is a fixed point, the fact that A *and -A* are neither true nor false with respect to e, means that neither *T "A"* and *T"-A"* will be true or false with respect to o" and therefore neither will their negations. This follows immediately from the fact that the fixed points satisfy the adequacy condition CA. Therefore (93) and (94) will *not* be true with respect to œş, as we would like to say. And this means that, although the adequacy of o allows us to say that L is able to express its notion of truth through the predicate *T,* and indirectly the notion of falsity, the same language does not have the resources to express arm *truth* and *nori falsity*. The only way to correctly describe the situation is to ascend to the me- tal language:

(93') non nt 1= *T "A"*
(94') non-re, 1= *T"-A"*

Well, this means that L is not perfectly closed semantically: at least in certain cases, the use of a more expressive metalanguage than L is necessary in order to be able to express semantic facts concerning L.

That things have to be this way is moreover evident if we return for a moment to the enhanced version of the liar, which in Italian leverages precisely one of the notions in question:

(95) This statement is not true.

If (95) is true, then the facts correspond to what it says, so it must not be true. On the other hand, if (95) is not true, then the facts do not correspond to what it says, so it must be true. In short, (95) is true if and only if it is not true: a contradiction. We had already noted that this version of the paradox is particularly insidious because it does not depend on the assumption of bivalence. Now, we know that (95) can be translated into L, for example through a formula such as (90). It is evident, therefore, that if we could also translate into L the reasoning we have just done, and in particular the assertion of the non-truth of (90), we would also find ourselves in L with a contradiction despite the semantics not being bivalent. The existence of a discrepancy between object language and meta-language, however contained, is therefore *unavoidable* penalty of inconsistency.

To what extent this result constitutes a serious limitation of the theory is still a matter of debate. It is certainly not the drastic limitation that characterises theories that impose a strict and absolute respect for the linguistic hierarchy for every use of the truth predicate, and on a practical level we can also say that it is an entirely irrelevant limitation. But the fact that *which* use*(s)* remain(s) illegitimate certainly presents itself as a considerable limitation on the theoretical level, and in this sense it can be assumed that the success of the theory is only partial. In the words of Kripke himself: 'The spectre of Tarski is still among us'.