JOHN MARTIN

A MANY-VALUED SEMANTICS FOR CATEGORY MISTAKES*

Zeugma may be informally defined as a figure of speech in which the single occurrence of some word occurs as part of several word groups and has a different meaning in each. Fowler cites some genuine examples found in literature:

> Half-clad stokers toasting in an atmosphere consisting of one part air to ten parts mixed perspiration, coal dust, and profanity.
> Such frying, such barbecueing, and everyone dripping in a flood of sin and gravy.

But logical theories of the semantics of such expressions divide on whether they are equivocal.[1]

The views that read zeugmas as equivocal are probably the best known. Fred Sommers expresses the basic idea as follows:[2]

> Each of the sentences 'The chair is hard' and 'The question is hard' is significant, yet 'The chair and questions were hard' is a category mistake.

Relative to a given situation two subject-predicate sentences could be bivalent, even true, yet their conjunction meaningless. In bivalent theories, two true conjuncts might be said to yield a false conjunction. In non-bivalent accounts, the two true conjuncts would yield a non-true conjunction. If either is right, classical logic is wrong in its truth-tables for the simple connectives, and truth-functionality is not universal.

Such readings of 'hard' in Sommers' example as equivocal are intriguing but dubious. They are intriguing because they recognize a new systematic elaboration of the rule that sense determines reference. Ordinarily, in any given world and context of utterance, the sense of an expression is thought to be constant throughout all sentences in which it occurs. But on this interpretation of zeugma, an expression's sense is determined relative to the sentence (or, perhaps, sentence token) in which it appears. It seems to be the reference of the subject term that helps determine the sense of the predicate. One attraction of this idea is that it seems to reflect some of the complexity of language.

Synthese **31** (1975) 63–83. *All Rights Reserved*

But this one idea will not suffice for a theory. Granted that a predicate may have different senses in two atomic sentences, it must still be decided which senses it will have in its conjunction. Presumably, it will be equivocal, its sense in any occurrence being that of its atomic part. But then the truth-value of the whole is determinate, even true, and there is no explanation for the oddity of the whole. It is possible that a satisfactory solution could be found to this problem and a semantics developed that would relativize sense to an expression's occurrence in a systematic way. Such a theory may prove viable, but we shall not pursue the question here.

The alternative approach reads the contained predicate as univocal and leads more naturally to the view that the compound sentence containing it is a category mistake. The predicate, on this view, has the same meaning in both conjuncts, but makes sense in only one. The other conjunct is a category mistake. From this point, the theory may develop in various ways depending on the semantical representation of category mistakes and of their effect on the truth-tables for the connectives. On most accounts, the resulting compound sentence often shares the same semantic deviance as its contained atomic part. We shall investigate each of these issues in turn: first, what are category mistakes and whether they constitute a significant part of linguistic behavior; second, how should they be semantically represented; and third, if they should be represented by a non-classical truth-value, what are its best principles of projection.

Category mistakes may be illustrated first by examples:

The GOP is deductible.
The barn is grammatical.
Earth is more honest than *C* sharp.

They may also be characterized in terms of language games. Classical logic proffers a picture of language in which every statement question must have a yes or no answer, in which no rejection of the question is countenanced. A simple paradigm of such linguistic behavior is the presentation language game. Situations in which an expression *E* applies to some things but not to others, as a predicate applies to the members of its extension, is represented in this game by a judge who on presentation of each object pronounces 'yes' or 'no'. But the description of this situation may be refined in such a way that classical logic appears simplistic. It abstracts away from a recognizable feature of the situation that amounts to category mistakes, a feature that a better theory would capture. A more accurate description would allow for the fact that some statement questions are never asked, and some propositions never asserted. Such is the case with the examples displayed above.

Perhaps no one has remarked on the generic feature of language encompassing category mistakes better than Wittgenstein. On the use of 'composite' he says,[3]

> The question "Is what you see composite?" makes good sense if it is already established what kind of complexity – that is, which particular use of the word – is in question. If it has been laid down that the visual image of a tree was to be called "composite" if one saw not just a single trunk, but also branches, then the question "Is the visual image of this tree simple or composite?", and the question "What are its simple component parts?", would have a clear sense – a clear use... Asking "Is this object composite?" *outside* a particular language-game is like what a boy once did, who had to say whether the verbs in a certain sentence were in the active or passive voice, and who racked his brains over the question whether the verb "to sleep" meant something active or passive.

One variety of this absurdity ought to be associated with category mistakes. They might be explained within Wittgenstein's framework as follows. Relative to any language game in which an expression E applies to some things but not to others, there may be expression pairs like 'yes' and 'no' designed to reflect whether E applies in particular cases. And in exactly the same situations in which these pairs may be operative, we can imagine attempts to apply E to some objects unprovided for by the language game, objects for which E does not have, in Wittgenstein's words, "a clear sense – a clear use". To the two part language game of affirmation and denial is added a third, that of rejection of the question. Thus, a category mistake may be informally characterized as a predication unprovided for by the rules of a presentational language game.

It will be instructive to construct one such language game in detail. Suppose that within a group of soldiers there is a language game operative such that upon polishing his shoes, buckels, metals, or various insignia, a soldier simultaneously presents the article to his sergeant and utters the word A. The sergeant, in turn, replies by uttering either the word A or the word B. If it is A, the soldier may quit polishing the article, but if it is B, he must go back to work. Translations of A and B into natural language spring readily to mind. It is also possible to imagine a soldier in this community presenting an inappropriate object, like a teapot, which is not the sort of thing to be polished in such contexts. It is contingencies such as these that are provided for in a formal theory by the recognition of category mistakes.

By this means too we can locate category mistakes within a wider conceptual frame. Notice that there may be many things wrong about a particular presentation than the 'type' of the object presented. If a soldier presents a garbage can to his buddy while they are both on K.P., saying *A*, there are more things wrong with the presentation than that the object presented is of the wrong sort. Each condition for the operation of a language game may be called a *presupposition*. Then, a category mistake is a failure of what we may call sortal presupposition.

On this account of category mistakes, the univocal reading of zeugma seems right on the point that one of the contained atomic sentences is a category mistake. It is perfectly fair to interpret 'the question is hard' as a category mistake given that by 'hard' we mean resistant to touch. A test would consist of substituting the appropriate synonym for 'hard' and judging whether the result is a category mistake, whether in ordinary usage it fails of sortal presupposition. It remains to be decided, however, how we are to develop the semantics of category mistakes.

One policy is to hold fast to classical logic. An interesting example of this alternative is suggested by remarks of Quine.[4] According to him, atomic sentences violating sense restrictions are false, and ambiguities of natural language may give way to distinct expressions in a formal language. Consider again the case of the hard question and chair. Either 'hard' is translated by a unique symbol, in which case one of the conjuncts is false, or it is translated by two symbols, one for each of its senses, in which case both conjuncts may be true. In the former case, zeugmatic expressions appear in the formal idiom and are treated by the bivalent rendering of meaningless sentences. In the latter, they are suppressed by convenient reformulation. It is a fair summary of the view, however, to say that when zeugmas appear in the formal language and are not banished by the introduction of new vocabulary, they possess a fixed sense and a meaningless conjunct. But unluckily for this theory, the step to many-valued semantics for category mistakes is irresistible.

As our discussion has shown, category mistake is a meaningful empirical concept. The distinction is there in the linguistic data to be drawn. A fully adequate theory would account for it. It is less clear, however, that it should be represented in a formal theory by non-classical truth-values, or that there is a third sentential property in addition to truth and falsehood to which these values might correspond. The issue is whether the traditional, pre-analytical concepts of truth and falsehood can be understood to admit a third possibility. An affirmative resolution may be grounded in our intuitions about the concept of presupposition. In one natural sense of this notion, it is connected to the concepts of truth and falsehood: in order for a sentence to be either true or false, its presuppositions must be satisfied. A simple example is grammatical correctness. Another example, existential presupposition, underlies the important developments in free logic. If semantics is to be rich enough to recognize presuppositions in this sense, truth and falsehood must move aside to admit another category of sentences. Notice that each of the examples we have just cited may be viewed as instances of the more general notion of presupposition we characterized earlier. Each is a condition for the appropriateness of a predication in a presentational language game. We may reasonably suppose, then, that a theory of sortal presupposition will be non-bivalent.

There is, however, a further hurdle that a many-valued semantics must cross that is neither empirical nor conceptual, but logical. Though classical semantics may be inadequate in some respects, classical logic, on the whole, is not. With the possible exception of the cases studied by relevance logic, it would be a great strain on our logical intuitions to reject any of those arguments countenanced as valid in classical logic. We will, in fact, adopt as a criterion of adequacy for a sortal theory that it not reject any classically valid arguments or theorems. Unfortunately, many-valued theories do not in general meet this criterion, and this seems to be true in particular of the three-valued matrix theories most frequently used in the analysis of category mistakes. We shall, however, explore these approaches in some detail in an effort to abstract from them intuitions or principles about category mistakes that we might then apply to other many-valued theories that accord better with classical logic.

Most truth-functional three-valued accounts make use of either the weak or strong connectives of S. C. Kleene. We shall call the corresponding matrices K_w and K_s. The tables for the usual connectives are as follows:

K_w

	$\neg$	&	$T\ F\ N$	$\vee$	$T\ F\ N$	$\rightarrow$	$T\ F\ N$
T	F		$T\ F\ N$		$T\ T\ N$		$T\ F\ N$
F	T		$F\ F\ N$		$T\ F\ N$		$T\ T\ N$
N	N		$N\ N\ N$		$N\ N\ N$		$N\ N\ N$

K_s

	$\neg$	&	$T\ F\ N$	$\vee$	$T\ F\ N$	$\rightarrow$	$T\ F\ N$
T	F		$T\ F\ N$		$T\ T\ T$		$T\ F\ N$
F	T		$F\ F\ F$		$T\ F\ N$		$T\ T\ T$
N	N		$N\ F\ N$		$T\ N\ N$		$T\ N\ N$

The former are also the tables for Bochvar's internal connectives, and the latter for a system of Łukasiewicz but with the conditional defined in terms of negation and conjunction.[5] Atomic sentences are thought to be made up of subject and predicate. To each predicate in a possible world is assigned a set of objects of which it is either true or false, which is variously called its category, sort, or significance range. The predicate's extension is then a subset of its category. An atomic sentence is assigned T, F, or N according to whether its subject falls in the predicate's extension, in its category but not in its extension, or outside its category. Values for the molecular sentences are then calculated according to the matrix in question.

Leonard Goddard in [5], pp. 251–255, has one such theory based on the weak connectives and extended to an analysis of zeugma. In addition to predicate categories and the weak connectives, he introduces a theory of ambiguity that relativizes sense to context. Once the senses of terms are fixed relative to a context of utterance, the sentences they make up may be evaluated relative to a world they describe. In different contexts in the same world, senses may differ, so sentences receive truth-values relative to both a context and a world. We may define a formal semantics for such a theory by identifying senses or intensions with functions from possible worlds to extensions, and relativizing senses to a set of indices called contexts. Let Syn be a propositional calculus syntax except that the atomic sentences are made up in the usual way by concatenating n proper names after a predicate of degree n. Let ξ range over the connectives and O_ξ over their corresponding operations in K_w.

DEFINITION *A K_w indexial language* with structure to its atomic sentences is any $\langle$Syn, WxC, $S\rangle$ such that W, $C \neq \phi$, W is a family of sets and S is a function on names t, predicates P^n and sentences A such that

(1) for any expression e, $S(e)$ is a function on WxC;

(2) $S_{cw}(t) \in w$;

(3) $S_{cw}(P^n)$ is a function from w^n (the nth Cartesian power of w) into $\{T, F, N\}$;

(4) if $A = P^n t_1 \ldots t_n$, then $S_{cw}(A) = S_{cw}(P^n)\,(S_{cw}(t.), \ldots, S_{cw}(t_n))$;

(5) if $A = \xi(A_1 \ldots A_n)$, then $S_{cw}(A) = O_\xi(S_{cw}(A_1), \ldots, S_{cw}(A_n))$.

Intuitively, if $x \in w \in W$, x is in the domain of w. Since the language has predicates we can define the ancillary notions of extension and category. Let the extension of P^n relative to world w and context c be $\{\langle x_1, \ldots, x_n\rangle : S_{cw}(P^n)\,(x_1, \ldots, x_n) = T\}$, and let the category or type of P^n relative to world w and context c, briefly $T_{cw}(P^n)$, be $\{\langle x_1, \ldots, x_n\rangle : S_{cw}(P^n)\,(x_1, \ldots, x_n) \in \{T, F\}\}$. Let $S_c(e)$ be the function f on W such that $f(w) = S_{cw}(e)$, and likewise for $T_c(e)$.

Look now at zeugma as it appears from this linguistic vantage point. All sentences including zeugmas must be evaluated as to truth only after the meanings of the words have been agreed upon and a world to be described has been chosen. What makes zeugma possible is the existence of predicates that under different senses have disjoint ranges of significance. For example, 'hard' meaning resistant to touch covers tables but not questions. The opposite holds when 'hard' means perplexing. If such a predicate is jointly said of objects from rival ranges, the result is a zeugma. Further, the root intuition is satisfied: zeugmatic compounds contain a meaningless part and are themselves meaningless.

This process of adapting the classical theory of zeugma to a three-valued significance theory has lead to several innovations which quite clearly improve on the old ways. Ambiguity is now possible with the desirable result that the formal language exhibits a property of natural language lacking in the classical theory suggested by the remarks of Quine. At the same time, by preserving truth-functionality, the notion of truth continues to have a straightforward analysis. But foremost among the advantages of this development of Goddard's idea is that the concept of sense or intension has been explicitly tied to semantic categories. The importance of this last step requires discussion.

The joining in one theory of both senses and categories is welcome from the viewpoint of scientific unity because it integrates intensional logic with significance theory. But more importantly from the philosophical point of view is that both senses and categories have been given representatives in a formal semantics in which they stand in clear relationships to each other. It is rather surprising that in the history of semantic categories very little has been said about the traditional concept of an expression's sense. In discussions of a sentence's meaningfulness, very little has been said of meaning. The reason is not that the traditional notion of meaning is unfamiliar. Meanings are the kind of entities, of dubious ontological status, shared by an expression and its translation, its *definiens* and its synonyms. It is also said to be the same as an expression's information or propositional content, the idea or thought it expresses, and it is said by some to be what is discussed in propositional attitude statements. In short, it is the notion of sense as it is studied in intensional logic and represented in the foregoing semantics by $S_c(A)$. A somewhat different concept of meaning is that which appears in the usual discussion of semantic categories. Meaning in this context assumes the form of a property of sentences called meaningfulness which sentences have when bivalent and lack otherwise. When not bivalent, sentences receive a third value that may be called meaninglessness. In the preceeding formal semantics this notion of meaning is represented by the trivalent assignments of truth-values relative to points of reference.

A few intuitive rules exist that relate these two concepts of meaning and which may serve as measures of the success of their formal representations in the same semantical system.

Sören Halldén, [6], p. 37, has observed:

> The etymology of the word 'meaningfulness' suggests that a meaningful entity is an entity which has 'meaning'. Some writers on epistemology have undoubtedly used the word thus. Now, it may of course be the case that there exists some connection between the property of being true or false, and the concept of meaning. However, that is something to be proved. It must be stressed that there exists no directly discernible connection between the two concepts.

Could not a sentence be neither true nor false and yet meaningful? The negative answer underlies the facile identification of bivalence with meaningfulness in significance theory and derives from examples like

(1) The number two is red

which seem neither bivalent nor comprehensible. But it is not difficult to drive a wedge between the two notions. Look at the example more closely. Having fixed on the natural language meanings of the subject and predicate, we find the sentence is neither true nor false and indeed that it is neither true nor false in any possible world in which we continue to use these words with their ordinary meaning. Compare this to another example that is only contingently absurd:

(2) Einstein's most important discovery supports combustion.

Even though relativity theory neither does nor does not support combustion, (2) is meaningful. Perhaps it is more accurate to say it is comprehensible to the extent that we would know how to investigate whether it is so or not, or whether it is neither.[6] We may tie this investigation into the truth-value of (2) with the notion of possible world as follows. We know the sentence is possibly true and may investigate whether the actual world satisfies it. Hence we shall say a sentence is meaningful if, and only if, in some possible world it is bivalent. By this way of speaking we preserve the intuitive meaninglessness of (1) and the meaningfulness of (2), and refined the identification of meaningfulness with bivalence.

Paradoxical sentences like

C is not true

in which *C* is the sentence displayed on line 20 of this page, are often said to be contingently absurd and intelligible at the same time. They fit neatly with our scheme.

Two additional principles relating categories and senses can also be used to evaluate their formal representation. Simply stated, they say that sense determines category but that the converse often fails. The first says that if A and B are synonymous, their semantical categories are identical. Suppose John is a bachelor, Jim is not and Jane neither is nor is not. It follows then that John is an unmarried man, Jim is not and Jane neither is nor is not. On the other hand, the category of 'bachelor' and 'married man' coincide but their senses differ. Therefore categories may be identical and senses not. The renderings of this rule in the terminology of the foregoing semantics are:

(a) for any L, $S_c(P)=S_c(Q)\rightarrow(\forall w)\ (T_{cw}(P)=T_{cw}(Q))$;

(b) for some L, c, w, P and Q, $S_c(P) \neq S_c(Q)$ and $T_{cw}(P) = T_{cw}(Q)$.

Both are easily verified.

There is, however, a major flaw in this and other three-valued theories based on K_w or K_s. It has proven very difficult to define within them a convincing semantic counterpart to classically valid argument. Indeed, most attempts reject some classically valid arguments. See for example Hans Herzberger [9], and J. N. Martin [15]. For this reason, we shall turn later to a four-valued theory designed to retain the best of the three-valued ones yet preserve classical logic. But first there is more to learn from the three-valued matrices. Both K_w and K_s embody principles of categorial relations that we should abstract and evaluate for their eventual help in constructing a new theory.

One sort of justification of these matrices is based on the assumption that people do, in fact, utter conjunctions and disjunctions containing category mistakes as atomic parts. The weak connectives may be defended by making N represent semantic deviance and observing that deviance behaves in principle analogously to purity or grammaticality: a flaw in a part is a flaw in the whole. The strong connectives may be defended on the same interpretation by the claim that it is sometimes possible to use a compound sentence to convey information when it has sortally defective atomic parts. Language, it would be argued, is primarily a tool, and an expression should be used if possible. If part of a conjunction is false, the whole should be rejected, and if part of a disjunction true, the whole accepted. Hence, in addition to holding that the classical truth-tables are right for the values they assign to the T and F, a trait K_s shares with K_w, the strong connectives recognize a further concession to classical logic. If a truth-value is dominant in the classical tables it remains so in K_s. If we expand the presentational language game to include conjunctive and disjunctive presentations – whatever they might be – we can represent these rival points of view as follows. In the soldier's game, for example, the sergeant according to the insight of the weak connectives should be a purist, and reject any presentation, however deviant, as inappropriate. According to those of the strong connectives, on the other hand, he should be more practical. Keeping his mind on the goal of gleaming brass and leather, he should not turn aside from a conjunctive presentation one part of which is inadequately polished; rather he should say the equivalent of, 'No, not yet. Get back to work.' In this way the soldier will be encouraged, perhaps by making further individual queries with the various conjuncts, to accomplish the goal of the language game. A disjunctive presentation with one adequately polished part should likewise be accepted. Unfortunately, the arguments for both positions are based on an assumption that is largely false, that people do, in fact, articulate category mistakes and use them in compound sentences.

Category mistakes are pointless, we know they are, and with the exception of infrequent zeugmas, we avoid them in compound sentences composed from 'and' and 'or'. Such is part of the linguistic wisdom of every language speaker. This fact does not disrupt the trichotomy among atomic sentences. There are still those presentations that would be accepted, those that would not be accepted, and those that though the situation is right are never preformed because they are unprovided for by the rules of the game. The sortally deviant atomic sentences, then, are conspicuous by their absence. So too are sortally deviant conjunctions, disjunctions, and conditionals. In this respect, which is perhaps the most important, the analogy between deviance and purity is upheld in its unqualified form. The difficulty that troubles the Łukasiewicz matrix is that it depends on the false hypothesis that people do in fact try to form conjunctions and disjunctions from sortally deviant parts. If they did, it might be right.

It is significant that most of Ryle's examples of category mistakes fall into one of two classes. They either involve language learners, or fully competent speakers who are ignorant of the proper category of the subject under discussion. But neither of these are convincing as counterexamples to the claim that, with the exception of zeugma, category mistakes are absent from ordinary speech. The first because this claim is tacitly restricted to the fully competent speakers, and the second because it is probably an empty set. Consider what it would be not to know the category of an object before you. Let us set aside philosophical language as perverted. We might be engaged in classification and be stumped about the species of something. But then to make our point, we must assume that natural kinds have something to do with semantic categories. If any natural kinds are semantic categories, then they would be probably only the broadest ones, those with which it would be hard to make classificatory errors, and in any case, the literature relating the two notions is regretably, but conspicuously, non-existent. Perhaps Aristotle was wrong to label the earth as non-celestial, or Jonah to call a whale a fish. But such cases, drawing on the hindsight of modern science, seem to reflect conceptual revolutions, alterations in the meanings of words, more than they do category mistakes. On the whole, then, we shall accept the principle underlying the weak connectives that an expression with a sortally deviant part is itself sortally deviant.

Thomason has recently advanced in [26] a three-valued sortal theory that avoids the major shortcoming of the matrix theories based on K_w and K_s. Its logic validates all the classically valid arguments. In addition, it appears to be capable of incorporating all the good features of its predecessors. There is, however, one exception. Its projection of truth-values may be described, with certain qualifications, as obeying the principles underlying the strong connectives. These qualifications amount to certain further concessions to classical logic. Before explaining these matters in detail, we should point out that his theory has an alternative intuitive basis grounded in the notion of supervaluation and is not put forward as an implementation of the ideas of K_s. To the extent, however, that K_s characterizes its projection and is itself inadequate, so too is the supervaluation theory. It should also be observed that Thomason is aware of the similarity of his theory to K_s. On the matrix disjunction he writes in [25], p. 21, "In many of its details [it] resembles the approach I wish to adopt, and so my reasons for avoiding this alternative depend more on global considerations than on detail".

Thomason's projection may be characterized in terms of K_s as follows. We shall use the definitions and notation of [26], except that we shall use as truth-values $\{0, 1/2, 1\}$ with 0 for false and 1 for true. Let VE be an arbitrary valuation on L relative to a sortal specification E on a logical space S. Let $A \Vdash_{VE} B$ iff $\forall f \in \mathrm{Biv}(VE), f(A)=1$ only if $f(B)=1$.

THEOREM[7] For any formulas A and B, $VE(A \to B) = \max[1 - VE(A), VE(B)]$ unless $VE(A) = VE(B) = 1/2$ and $A \Vdash_{VE} B$, in which case $VE(A \to B) = 1$.

Similar results obtain for conjunction and disjunction.

Since A classically entails B only if $A \Vdash_{VE} B$, the theorem records a concession to classical logic. It is similar to that made by Łukasiewicz who advanced a table for implication like that of K_s except that $\langle 1/2, 1/2 \rangle$ is taken into 1, rendering the classically valid $A \rightarrow A$ a tautology in the new matrix. Being unconstrained by the matrix form, supervaluations are able to defer to classical logic in all such cases.

As an alternative to Thomason's proposal and an improvement on the matrix theories, we will now advance a four-valued explanation of category mistakes developed within the general theory of presupposition advanced by Herzberger in [8]. Applying Herzberger's theory to the problems of sorts as Thomason did van Fraassen's is particularly appropriate since Herzberger's account was originally intended to serve as an alternative to van Fraassen's general supervaluation account of presupposition. The success of our version of sortal presupposition will serve as an interesting measure of the strengths of the two general theories.

Herzberger's theory is the rigorous development of a philosophical analysis of the concept of truth. Its main tenet is that to say of a sentence that it is true means that it has two properties. The first is that it corresponds to the world. The second, which is omitted in the usual theory of truth, is that all the presuppositions of the sentence are themselves true. This property we shall call *presuppositional security*. A sentence is true, on this view, if and only if it both corresponds and is presuppositionally secure. It is precisely this proposition that we argued was intuitively sound earlier in this essay.

A second tenet of the theory is that logical inference is a function not of truth, but of correspondence. The argument from A to B is logically valid if and only if whenever A corresponds, so does B. The divorce of logic from truth is exciting. Though it deviates from a long tradition, it is defensible according to principles nearly as venerable, for truth has usually been identified with correspondence. It is this simple equation that is called into question. Classical conceptions of truth have succeeded only at the expense of ignoring presupposition. In elementary logic, for example, we could teach that 'the sky is blue' and 'the sky is not colored' might both be true. To avoid these nonacceptable consequences, we should be willing to admit presupposition as a factor of truth: a sentence is true if, and only if, it both corresponds and is presuppositionally secure. We thereby rechart the region of Platonic Heaven in the neighborhood of Truth. The result, as it stands, is a plausible picture. It deals fairly with traditional views and transcends them in its scope by explaining the role of presupposition. It is an adequate basis for a formal theory.

The formalization of these intuitions requires that sentences be assigned a two-dimensional truth-value, i.e., an ordered pair. Values on the first dimension stand for the property of correspondence and those on the second dimension for presuppositional security. In any situation, those sentences that correspond receive 1 on the first dimension, and those that do not, 0. Those that are presuppositionally secure receive 1 on the second dimension, and those that are not, 0. Those that correspond and are secure are according to the theory true, and they receive the truth-value $T=\langle 1, 1\rangle$. Those that fail to correspond but are secure receive $F=\langle 0, 1\rangle$. Those that fail of presupposition but correspond or do not, receive $t=\langle 1, 0\rangle$ and $f=\langle 0, 0\rangle$, respectively. In this fashion, a four-valued valuation may be constructed from two two-valued ones. Let υ and v be two-valued valuations, the characteristic functions of the set of sentences that correspond and are presuppositionally secure, respectively. Then a four-valued valuation w may be constructed as follows: $w(A)=\langle \upsilon(A), v(A)\rangle$.

Logic is, according to the theory, a function of correspondence and is to remain classical. It is therefore required that each correspondence valuation assign its values according to the matrix for classical logic, call it C, and be a member of the set of all classical valuations $\mathscr{V}\!al$. The set of presuppositional security valuations Val is for the moment left unspecified. Wal, the set of four-valued valuations, is constructed as explained from $\mathscr{V}\!al$ and Val. The set of designated values for our four-valued language is $\{T, t\}$, and it follows that the semantic entailment relation for L is perfectly classical. Let *PCS* be the usual syntax for the propositional calculus.

Definition. A *two-dimensional* language is any $\langle$PCS, Wal$\rangle$ such that, for some set Val of bivalent PCS valuations (possibly non-classical) and $\mathscr{V}\!al$, the set of classical valuations generated by C,

> Wal $=\{w$: w is a function of the sentences of PCS and there is a $\upsilon\in\mathscr{V}\!al$ and $v\in$Val such that for any A, $w(A)=\langle \upsilon(A), v(A)\rangle\}$.

T and F continue to have the same interpretation they have in the three-valued matrix theories, and t and f correspond to a division of the sentences we there called N. Ideally, we should be able to explain the difference between t and f. Possibly the most straight forward way to do so is by means of counter-factuals. A sentence A is t means that if all the presuppositions of A were met then A would be true; it is f means that in the same circumstances A would be false. Unfortunately, it is rather hard to get any solid intuitions about such hypotheticals because it is rather hard to imagine a situation in which sortal errors have their presuppositions satisfied. Should 'Socrates is legible' be t of f?

These truth-values should correspond to some distinction in the informal conceptual theory. Their interpretation is an important bridge between the technical constructions and traditional issues, and contains much of the theory's philosophical purport. In general, the primary justification for explanatory tools like truth-values is that they represent properties of sentences that are familiar – if not completely understood – from the history of philosophy and less formal accounts of language. But what are the properties captured by the various truth-values? One, surely, will be truth, for the theory claims to be semantic. As for the others, it is good to bear in mind the ideal that every distinction in the theory should represent a real difference. Every theoretical notion should be a more precise version of a pre-analytic one, and the web of technical concepts should accurately reflect at least the uncontroversial parts of previous usage. This interpretation of truth-values is part of what Thomason calls the "theoretical criteria of motivation" underlying a formal theory. On these criteria he writes in [25], p. 14, "They require us to give a plausible reason for each decision in constructing the theory; the finished product should be made to seem inevitable."

One way to explain the distinction between t and f is as an abstraction from the presentation language game. Recall for the moment the language game of the polishing soldiers. Suppose a soldier were to present an object which is well polished but not provided for in the language game, for example, a shiny silver teapot. We might characterize such situations by t. Likewise, we may characterize the presentation of tarnished object unprovided for by the rules of the game by f. By characterizing such situations by t and f, and proper and improper presentation within the rules by T and F, all four truth-values are interpreted. In general, some objects are such that the rules of a presentational language game might be expanded to admit them as appropriate. Such objects that when admitted and presents were acceptable would be assigned t, those that were not, f. The teapot, for example, could be added to the soldiers' language game. If it were shiny, it would get *t*, if tarnished, *f*. Hence, there is, in fact, a distinction in linguistic behavior between types of nonsense corresponding to *t* and *f*, and our formal theory may proceed understood as abstracting from a real distinction.

Though we have said enough to justify our use of *t* and *f*, there is more to be said about the scope of the distinction. For, though it is acceptable to include under *N* any object not provided for in the presentational context, not all such objects clearly fall under *t* or *f*. While it is true that the rules of a language game may be relaxed to incorporate new objects, they cannot embrace the universe. A soldier might present to his sergeant a song, but the sergeant could not evaluate this object as to its degree of luster. Some objects are not the sort of things that are shiny in any context. It is among these incorrigibles that there does not seem to be a real difference between *t* and *f*.[8] To the extent that this distinction must be seen as arbitrary, as, perhaps, a technical convenience, we fall short of a fully articulated formal theory.

But being completely interpretable over the whole range of linguistic behavior is a goal that probably no formal theory has yet met, and it is only one measure of success. We want to know also the long range conconsequences of the formal theory, whether it also sheds light on additional problems and poses interesting new questions. One question it is fruitful to pursue within the context of a product language is that of the matrix behavior of the presuppositional dimension. Whether the presuppositions of *A* are met is determined by both dimensions. Its presuppositions must both correspond and have their own presuppositions met, but notwithstanding this dependence of one dimension on the other, we may still pursue the question of how the presupposition-values of the parts effect that of the whole. If, in fact, there is a functional correlation between part and whole on the presupposition dimension, then presupposition-values conform to a matrix. But whether presupposition is functional is still, as far as our investigation is concerned, an open question. Its philosophical importance lies in the fact that various projections of presupposition-values will represent philosophical policy. We must answer the question of which matrix is relevant, if any, according to principles and intuitions about presupposition. Those principles and insights germane to the issues are ones we have not yet discussed from within the four-valued framework. They concern the relation of presupposition to the sentential connectives.

One rule underlying K_w is, recall, that if any of an expression's parts fail of presupposition, so does the whole. Herzberger in [8], Section VII, has applied this principle to a two-dimensional language by means of the following tables for the presuppositional dimension:

	$\neg$	&	1	0	$\vee$	1	0
1	1		1	0		1	0
0	0		0	0		0	0

Call this matrix K_w^2. A compound has its presuppositions met just in case all its parts do. For any valuation v for this matrix, $v(A \,\&\, B)=1$ iff $v(A \vee B)=1$ iff $v(A)=v(B)=1$, and $v(\neg A)=1$ iff $v(A)=1$. Hence, the parts of a compound do not necessarily number among its presuppositions. For it is possible for a whole to be true, to both correspond and be presuppositionally secure, yet its parts fail to be true. All we know from the fact that $A \vee B$ is true is that the presuppositions of A and B are met and that one or the other corresponds.

We may now define a four-valued matrix CxK_w^2, with operations χ_i, as the product of the classical matrix C, with operations ϕ_i with the matrix K_w^2, with operations ψ_i:

$$\chi_\neg(w, x) = \langle \phi_\neg(w), \psi_\neg(x)\rangle$$
$$\chi_\&(\langle w, x\rangle \langle y, z\rangle) = \langle \phi_\&(w, y), \psi_\&(x, z)\rangle$$
$$\chi_\vee(\langle w, x\rangle \langle y, z\rangle) = \langle \phi_\vee(w, y), \psi_\vee(x, z)\rangle$$

The tables for $\neg$ and & for the new matrix are as follows:

	$\neg$	&	T	F	t	f
T	F		T	F	t	f
F	T		F	F	f	f
t	f		t	f	t	f
f	t		f	f	f	f

The rules characterizing K_s are the same as those of K_w but with the following addendum: if the truth-value of a part determines that of the whole in the classical matrix, it should continue to do so combined with non-classical truth-values. If, for example, one conjunct is false, then the whole is, regardless of whether the other conjuncts are presuppositionally secure. In a two-dimensional framework, this means that the presupposition-values of a connective should be a function of both the correspondence and presupposition-values of the parts. Herzberger captures these ideas in the two-dimensional frame by calculating presupposition-values by means of functions that have truth-values as arguments. The presuppositional function for conjunction would be:

$\beta_\&$	T	F	t	f
T	1	1	0	0
F	1	1	1	1
t	0	1	0	0
f	0	1	0	0

The four-valued operation for conjunction would then be:

$$\chi'_\& (\langle w, x\rangle \langle y, z\rangle) = \langle \phi_\& (w, y), \beta_\& (\langle w, x\rangle, \langle y, z\rangle)\rangle$$

Together with X this yields a new matrix CxK_s^2. The table for $\chi'_\&$ is:

$\&$	T	F	t	f
T	T	F	t	f
F	F	F	F	F
t	t	F	t	f
f	f	F	f	f

Note that, for either four-valued matrix, if T and t are designated, the logic is classical, and if t and f are conflated, the resulting matrices are K_w and K_s. We may view three-valued theories as essentially two-dimensional ideas forced into a three-valued framework. By stepping up to four values, we get a clearer picture of our basic principles and at the same time preserve classical logic. We have shown, then, that two of Herzberger's four-valued systems have interpretations as sortal languages. It remains for us to choose the better of these and by combining it with the other good features of three-valued sortal semantics, to produce a synthesis that will avoid the shortcomings of previous accounts.

The language we shall define will be a classical two-dimensional language for CxK_w^2. Further, it will be the four-valued analogue of a K_w indexical language. We shall also extend the theory in a first-order direction by reading the quantifier as a sort of infinite conjunction. Fortunately, there are straightforward infinitary analogues to conjunction in both C and K_w^2. For perspicuity we adopt the substitution interpretation of the quantifiers. Let Zy/x be the result of replacing all the occurrences of x by y in Z. Then, let QCS be like Syn except that if A is a sentence, v a variable that does not occur in A, and t a name, then $(\forall v)$ (Av/t) is also a sentence. As before let ϕ_i range over the operations of C and ψ_i over those of K_w^2.

DEFINITION. *A two-dimensional sortal language L* is any $\langle QCS, WxC, S\rangle$ such that $W, C \neq \phi$, W is a family of sets, and S is a function on names t, predicates P^n, and sentences A such that

(1) For any expression in Domain (S), $S(e)$ is a function on WxC;
(2) $S_{cw}(t) \in w$, and $\forall x \in w \exists t\ S_{cw}(t) = x$;
(3) $S_{cw}(P^n)$ is a function from w^n into $\{T, F, t, f\}$;
(4) if $A = P^n t_1 \ldots t_n$, then $S_{cw}(A) = S_{cw}(P^n)\ (S_{cw}(t_1), \ldots, S_{cw}(t_n))$;
(5) if $A = \neg B$ and $S_{cw}(B) = \langle x, y\rangle$, then $S_{cw}(A) = \langle \phi_\neg(x), \psi_\neg(y)\rangle$;
(6) if $A = B\ \&\ C$, $S_{cw}(B) = \langle w, x\rangle$ and $S_{cw}(C) = \langle y, z\rangle$, then $S_{cw}(A) = \langle \phi_\&(w, y), \psi_\&(x, z)\rangle$;
(7) if $A = (\forall v)\ (Bv/t)$ and x, y are the least values such that $(\exists t')\ (S_{cw}(Bt'/t) = \langle x, y\rangle)$, then $S_{cw}(A) = \langle x, y\rangle$.

Let $D = \{T, t\}$. Then, *X analytically entails A in L* iff

$$\bigcap \{w\colon S_{cw}(B) \notin D\} \subseteq \{w\colon S_{cw}(A) \in D\}.$$

Further, *X logically entails A* iff for all L, X analytically entails A in L. Clearly, logical entailment is perfectly classical. The category of P^n relative to c and w continues to be $\{\langle x_1, \ldots, x_n\rangle\colon S_{cw}(P^n)\ (x_1, \ldots, x_n) \in \{T, F\}\}$.

University of Cincinnati

NOTES

* The ideas in this article were developed as part of my doctoral dissertation *Sortal Presupposition* (University of Toronto, 1973). I should like to acknowledge my debt to Professors Bas C. van Fraassen, my supervisor, and Hans G. Herzberger, my adviser, for their helpful and thorough criticism.

1 Though Noam Chomsky and others have defended the idea of a syntactic theory of category mistakes, which could be used in an explication of zeugma, Richmond

Thomason has pointed out that unlike grammatical correctness, sortal correctness is a function of context of utterance. Many of our subsequent remarks elaborate this view point. Compare Chomsky [3], Donald Hillman [10], and Thomason [2].

2 Sommers [24], p. 340. Compare Gilbert Ryle [22], p. 23, and Leonard Goddard [5], pp. 251–2. Sommers' positive theory consists of interpreting zeugma rather implausibly not as a compound sentence made up of atomic parts, but as a sort of quantified formula.

3 Wittgenstein [28], p. 22e. Of the accounts of category mistakes in the philosophical literature, our view is most similar to Bernard Harrison's in [7]. For a general account of formal semantics as abstraction from language games see Bas van Fraassen [27], pp. 1–6.

4 W. V. Quine [20], §§ 27, 33 and 47 (p. 229). Goddard in [5], p. 252 remarks that Quine denies zeugmas exhibit category mistakes. A more accurate description is that since Quine treats category mistakes bivalently, he also treats category mistakes in zeugmas bivalently.

5 On K_w see S. C. Kleene [12] and [13], p. 334, and compare J. Łukasiewicz [14]. Theorists who have used this matrix to analyse category mistakes include L. Åqvist [1], K. Donnellan [4], and R. L. Martin [16]. On K_s see Kleene [13], pp. 334–5, and D. A. Bochvar [2]. Applications to categories appear in S. Halldén [6], R. L. Martin [16], R. Routley [21], and K. Segerberg [23].

6 Such a criterion of meaningfulness is reminiscent of Quine's remarks on the alleged non-statementhood of logical impossibilities in [19], p. 202.

7 Proof is not difficult. For a development of the relatively simple theory required, consult my dissertation. Herzberger [9], pp. 28–29, and L. Karttunen [11], p. 186, remark on the similarity of supervaluation theory in general to the tables of K_s.

8 The distinction between corrigible and incorrigible presentations is similar to one of Bernard Harrison's in [7], pp. 315–318, between category mistakes that can be made meaningful by extending the meanings of terms and those which cannot. Herzberger [8], Section IX, proposes a completely universal distinction between t and f but at the cost of distinguishing between positive and negative elementary predicates.

BIBLIOGRAPHY

[1] Åqvist, Lennart, 'Reflections on the Logic of Nonsense', *Theoria* **28** (1962), 138–157.

[2] Bochvar, D. A., 'On a 3-Valued Logical Calculus and its Applications to the Analysis of Contradictions', *Matematiceskij sbornik* **4** (1939), 287–308.

[3] Chomsky, Noam, 'Some Methodological Remarks on Generative Grammar', *Word* **17** (1961), 219–241.

[4] Donnellan, Keith S., 'Categories, Negation, and the Liar Paradox', in [17], 113–120.

[5] Goddard, Leonard, 'Towards a Logic of Significance', *Notre Dame Journal of Formal Logic* **9** (1968), 233–264.

[6] Halldén, Sören, *The Logic of Nonsense*, A.-B. Lundequistska Bokhandeln, Uppsala, 1949.

[7] Harrison, Bernard, 'Category Mistakes and Rules of Language', *Mind* **74** (1965), 309–325.

[8] Herzberger, Hans G., 'Dimensions of Truth', *Journal of Philosophical Logic* **2** (1973), 535–556,

[9] Herzberger, Hans G., 'Truth and Modality in Semantically Closed Languages', in [17], 25–46.
[10] Hillman, Donald J., 'On Grammars and Category Mistakes', *Mind* **72** (1963), 224–234.
[11] Karttunen, Lauri, 'Presuppositions of Compound Sentences', *Linguistic Inquiry* **4** (1973), 169–193.
[12] Kleene, Stephen Cole, 'On a Notation for Ordinal Numbers', *Journal of Symbolic Logic* **3** (1938), 150–155.
[13] Kleene, Stephen Cole, *Introduction to Metamathematics*, North-Holland, Amsterdam, 1959.
[14] Łukasiewicz, Jan, 'On Three-Valued Logic', in [18], 16–18.
[15] Martin, John N., 'A Syntactic Characterization of Kleene's Strong Connectives with Two Designated Values', forthcoming in *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*.
[16] Martin, Robert L., 'A Category Solution to the Liar', in [17], 91–112.
[17] Martin, Robert L. (ed.), *The Paradox of the Liar*, Yale University Press, New Haven, 1970.
[18] McCall, Storrs, (ed.), *Polish Logic, 1920–1939*, Clarendon Press, Oxford, 1967.
[19] Quine, W. V. O., *Methods of Logic*, Holt, Rinehart and Winston, N.Y., 1959.
[20] Quine, W. V. O., *Word and Object*, M.I.T. Press, Cambridge, 1960.
[21] Routley, R., 'On a Significance Theory', *Australasian Journal of Philosophy* **44** (1966), 172–209.
[22] Ryle, Gilbert, *The Concept of Mind*, Penguin, Harmondsworth, Middlesex, 1969.
[23] Segerberg, Krister, 'A Contribution to Nonsense-Logics', *Theoria* **31** (1965), 199–217.
[24] Sommers, Fred, 'Types and Ontology', *Philosophical Review* **83** (1964), 522–527.
[25] Thomason, Richmond H., 'Semantics, Pragmatics, Conversation, and Presupposition', Unabridged manuscript, 1973.
[26] Thomason, Richmond H., 'A Semantic Theory of Sortal Incorrectness', *Journal of Philosophical Logic* **1** (1972), 209–258.
[27] van Fraassen, Bas C., *Formal Semantics and Logic*, Macmillan, N.Y., 1971.
[28] Wittgenstein, Ludwig, *Philosophical Investigations*, 2nd. ed., (trans. by G. E. M. Anscombe), Macmillan, N.Y., 1958.