What is Tarski's Theory of Truth?

Gila Sher

1. Interpretation as generalization

In "On the Concept of Truth in Formalized Languages", Tarski (1933) describes his project as follows:

> For an extensive group of formalized languages it is possible to give a method by which a correct definition of truth can be constructed for each of them. The general abstract description of this method and of the languages to which it is applicable would be troublesome and not at all perspicuous. I prefer therefore to introduce the reader to this method in another way. I shall construct a definition of this kind in connection with a particular concrete language and show some of its most important consequences. The indications which I shall then give in §4 of this article will, I hope, be sufficient to show how the method illustrated by this example can be applied to other languages of similar logical construction. (Pp. 167–168)

Tarski conceived of his theory as a *general method* for defining truth for a broad, if well defined, range of languages, but he chose to expound it through a single, simple example. This example, however, does not uniquely determine his general method, and the question arises as to how to generalize Tarski's example. Tarski clarified one aspect of this question, namely, how to extend his example to languages with indefinitely high order of variables, but many other fundamental issues were not addressed either in his original (1933) paper, or, indeed, in his later (informal) papers (1944 and 1969). The fact that Tarski did not address these questions is, of course, indicative of his attitude: Tarski was either unaware of these questions, or uninterested, or believed the answers were obvious and no further explanation was required. Today, however, the philosophical discussion has veered away from the technical matters that occupied Tarski in the 30's (partly, no doubt, due to his own thorough and successful treatment of these matters), and differences in attitude towards Tarski's theory are often grounded in differences in answers to the open questions. Even general attitudes towards the theory of truth (e.g., towards the possibility of a substantive, non-deflationist theory of truth) can be traced to implicit generalizations of Tarski's example.

In this paper I will study Tarski's theory through a few of its open questions and some of its generalizations. I will concentrate on the "reductionist approach" to Tarski's theory, exemplified by two generalizations due to Field. My critical investigation of these generalizations will not be directed at their exegetical virtues; rather, I will be interested in their viability as philosophical theories and in some of the challenges they face. I will begin with a brief introduction to the original goals of Tarski's theory.[1]

2. Aims of theory

We can distinguish three aims of Tarski's theory: a *philosophical aim*, a *methodological aim* and a *logical aim*.

1. *The Philosophical Aim*.[2] Tarski described his goal in constructing a theory of truth as philosophical in nature. The goal is to construct a materially accurate and formally consistent definition of the classical notion of truth:

> The present article is almost wholly devoted to a single problem – *the definition of truth*. Its task is to construct . . . *a materially adequate and formally correct definition of the term 'true sentence'*. This problem . . . belongs to the classical questions of philosophy . . . . [Ibid., p. 152. See also pp. 266–267]

By the 'classical question' of truth Tarski means the question of how to define the "classical", correspondence notion of truth:

> [T]hroughout this work I shall be concerned exclusively with grasping the intentions which are contained in the so-called *classical* conception truth ('true – corresponding with reality') . . . . [Ibid., p. 153]

Topoi **18**: 149–166, 1999.

The task of constructing an adequate theory of the classical notion of truth is, however, fraught with difficulties:

> This problem . . . raises considerable difficulties. For although the meaning of the term 'true sentence' in colloquial language seems to be quite clear and intelligible, all attempts to define this meaning more precisely have hitherto been fruitless, and many investigations in which this term has been used and which started with apparently evident premisses have often led to paradoxes and antinomies . . . . (Ibid., p. 152)

Tarski divides the philosophical task into two sub-tasks: (i) the *material task* of capturing the exact content of the correspondence notion, and (ii) the *formal task* of complying with the most rigorous standards of logical consistency and correct definition. But Tarski's treatment of these two tasks is not equal. In executing the formal task Tarski offers a substantive, in-depth analysis of the semantic paradoxes and a substantive proposal for preventing their occurrence, but in carrying out the material task Tarski offers no deep analysis of the correspondence notion of truth or the philosophical problems it gives rise to. The correspondence notion is treated either as well understood:

> the meaning of the term 'true sentence' in colloquial language seems to be quite clear and intelligible (ibid.),

or as a notion whose analysis is to be given elsewhere.

> A thorough analysis of the meaning current in everyday life of the term 'true' is not intended here. Every reader possesses in greater or less degree an intuitive knowledge of the concept of truth and he can find detailed discussions on it in works on the theory of knowledge. (Ibid., p. 153)

And the task is conceived as defining the bare skeleton of the philosophical notion of truth in a precise and formally correct manner, not as providing a new, deeper analysis of the material content of this notion. In this way the material task itself is construed as a formal task.

This situation creates special difficulties for a philosophical study of Tarski's theory. Not only does Tarski's reliance on a specific example leave the precise nature of his general method an open question, but many philosophical issues pertinent to his philosophical goal are not decided by this method. Not surprisingly, the interpretations we will discuss in the present paper involve revision and/or extension of Tarski's original theory.

2. *The Methodological Aim*. A secondary yet important goal of Tarski's theory is to contribute to the *methodology of the deductive sciences*, or (using Hilbert's terminology) *metamathematics*. Although by 'deductive sciences' Tarski primarily understood *mathematical* disciplines presented 'in the shape of formalized deductive theories' (Tarski, 1936b, p. 409), most philosophical interpretations of his theory have concentrated on its applications to science and everyday discourse. Tarski's intended contribution to the methodology of the deductive sciences was both negative and positive. His negative goal is described by Vaught as follows:

> [During the 1920s] Tarski has become dissatisfied with the notion of truth as it was being used. Since the notion "σ is true in $\mathfrak{U}$" is highly intuitive (and perfectly clear for any definite σ), it had been possible to go even as far as the completeness theorem by treating truth (consciously or unconsciously) essentially as an undefined notion – one with many obvious properties. . . . But no one had made an analysis of truth, not even of exactly what is involved in treating it in the way just mentioned. At a time when it was quite well understood that 'all of mathematics' could be done, say, in ZF, with only the primitive notion ϵ, this meant that the theory of models (and hence much of metalogic) was indeed not part of mathematics. It seems clear that this whole state of affairs was bound to cause a lack of sure-footedness in metalogic. . . . [Tarski's] major contribution was to show that the notion "σ is true in $\mathfrak{A}$" can simply be *defined* inside of ordinary mathematics, for example, in ZF. (Vaught, 1974, p. 161)

The positive goal contains both (i) the definition of central methodological (metamathematical) concepts, and (ii) the establishment of methodological (metamathematical) results. Among the methodological definitions and results that Tarski arrived at in the course of his work on truth (late '20s and early '30s) are the definition of *definability* (the notion of 'object (set) X is *definable* by expression *e*'), and the *undefinability result* (which says that the set of all true sentences of a reasonably rich, bivalent, deductive system S cannot be defined in S). (Tarski, 1931 and 1933, respectively.)

In a later paper Tarski mentioned another methodological goal: 'bring [the] method [of truth] into harmony with the postulates of the unity of science and of physicalism'. (Tarski, 1936a, p. 406) 'Unity of science' and 'physicalism' refer to methodological principles advocated by the Vienna Circle, and, as in the case of 'truth', Tarski translated these *material principles* into essentially *formal constraints*: (i) the definition of truth shall satisfy the requirement of formal rigor, and (ii) the definition of truth shall eliminate all semantic notions (since semantic notions are neither

logical nor physical). Not surprisingly, one of the differences between Field and Tarski is that Field's treatment of the physicalistic constraint is material rather than formal.

3. *The Logical Aim*. Indirectly, a theorist's aims are captured by what he/she does with his/her theory. And the most important uses that Tarski made of his theory of truth are unquestionably in logic: the semantic definition of 'logical consequence' and related meta-logical concepts, the establishment of model-theory (logical semantics) as one of the two main branches of meta-logic, and the numerous results Tarski established in this field. (See Tarski, 1933, 1936b, 1954, 1955; Tarski and Vaught, 1957, etc.)

In a sense, the logical aim is subsumed under the general methodological aim, but the logical aim raises issues that the general methodological aim does not. Briefly, the point is this: Not just any theory of truth can serve as a basis for *logical semantics*; only a theory which specifies the special contribution of *logical structure* to truth can. Today we usually take it for granted that an adequate theory of truth must specify the logical contribution. But what is the ground for this assumption? This assumption is natural *if* we assume that the theory of truth ought to have ramifications for logic. But aside from our familiarity with the consequences of Tarski's theory and our habit of thinking of Tarski's theory as *the* theory of truth, what is the basis for the latter assumption? The questions 'What is the exact status of logical structure in the theory of truth?' 'Must any adequate definition of truth assign a privileged status to logical structure?', 'How can a general philosophical notion serve as a basis for a very specialized branch of mathematical logic?' etc., are not fully answered by Tarski's exposition. Neither do these questions receive a full explanation by the existent generalizations of Tarski's theory (e.g., Field, 1972; Davidson, 1965, 1967a, etc.).[3] Field's generalizations, as we shall see below, are geared more towards Tarski's philosophical and methodological aims than towards his logical aim. The logical aim will be only obliquely discussed in the remainder of this paper. (In Sher (1998–99) I propose a new generalization of Tarski's theory which offers a new, full explanation of its connection to logic.)

3. Tarski's example[4]

Background Constraints (General Method). Tarski's method sets three general constraints on a definition of truth:

1. *A Formal Correctness Constraint*
   (a) Each definition of truth shall apply to a so-called *object language*, L, such that:
      (i) L is the formal language of some deductive science, S. (We will sometimes refer to L as 'LS'.)
      (ii) L does not contain its own semantic predicates.
   (b) The definition of truth for L shall be given in a *meta-language*, ML (or rather a meta-theory, MS), essentially more powerful than L(S).
   (c) ML (MS) is subject to the usual requirements of formal rigor and consistency.

2. *A Material Adequacy Constraint*

   Criterion T: MS shall include as theorems all T-sentences for L, i.e., all sentences of the form '$\overline{\overline{s}}$ is true iff $\overline{s}$', where '$\overline{\overline{s}}$' stands for an ML name of an L-sentence, *s*, and '$\overline{s}$' stands for an ML sentence which says the same thing as (has the same meaning as, constitutes a correct translation of) *s*.

   Comment: A colloquial example of a T-sentence is: ''Snow is white' is true iff snow is white'. This sentence captures the philosophical idea of truth as correspondence by reducing the truth predication 'TRUE ('Snow is white')' to the objectual ("worldly") predication 'WHITE (snow)'.

3. *A Methodological Constraint*
   The right hand side (the definiens) of a T-sentence shall not contain any semantic terms.

Based on these constraints Tarski provides the following guidelines for the construction of a meta-language, ML, of L, and of a meta-theory, MS, of S:

1. ML shall contain three types of vocabulary:
   (a) general logical vocabulary (including set-, or class-theoretic vocabulary);
   (b) structural-descriptive names of all expressions of L;
   (c) vocabulary equivalent in meaning to the vocabulary of L.

2. MS shall consist of three theories:
   (a) a logical theory (which includes sections of set theory or class theory, i.e., higher-order logic), more powerful than the logical segment of S;
   (b) a theory of the morphology of L;
   (c) a theory whose theorems are equivalent in content to those of S.

Comment: The notion of sameness of meaning was left unexplained; Tarski used the presumably clearer notion of translation to clarify his intentions, but no standard of correct translation was offered.

Notation: I will mark the two expressions correlated with an L-expression *e* – its ML-name and its ML-translation – by '$\overline{\overline{e}}$' and '$\overline{e}$', respectively. In the case of the left and right parentheses of L, however, I will use them as their own ML names and translations.

In addition to these guidelines Tarski introduces two *technical* directives:

1. The definition of truth for L shall be recursive.
2. The definition of truth for L shall be given in terms of the semantic notion of satisfaction.

These directives are intended to solve two technical problems:

(a) The languages of the deductive sciences contain infinitely many sentences;
(b) Sentences are generated from formulas rather than from sentences.

In carrying out these directives, Tarski, in fact, defines satisfaction recursively, and then defines truth directly (non-recursively) in terms of satisfaction.

The Particular Example. Tarski described his method of constructing a definition of truth for an arbitrary formalized language L by means of the following example:

Definition of Truth for the Language LC of the Calculus of Classes, C

I. *Vocabulary of LC*
   (a) primitive constants: '$\sim$', '$\vee$', '$\forall$', '$\subseteq$', understood as 'not', 'or', 'for all', 'is included in', respectively, where the last relation applies to classes of individuals;
   (b) variables: 'x_1', 'x_2', . . ., ranging over classes of individuals;
   (c) punctuation marks: '(',')'.

II. *Inductive Definition of 'Well-Formed Formula (or Wff of LC)'*[5]
   Let 'i' and 'j' be variables ranging over positive integers, and let 'Φ' and 'Ψ' be schematic letters representing arbitrary expressions of LC.
   1. $\overline{\forall}$i, j[$\ulcorner \overline{\overline{x}}_i \overline{\overline{\subseteq}} \overline{\overline{x}}_j \urcorner$ is a wff].
   2. Φ is a wff $\overline{\supset}$ $\ulcorner (\overline{\overline{\sim}}\ \Phi) \urcorner$ is a wff.
   3. [Φ is a wff $\overline{\&}$ Ψ is a wff] $\overline{\supset}$ $\ulcorner (\Phi\ \overline{\overline{\vee}}\ \Psi) \urcorner$ is a wff.
   4. Φ is a wff $\overline{\supset}$ $\overline{\forall}$i [$\ulcorner (\overline{\overline{\forall}}\ \overline{\overline{x}}_i\ \Phi) \urcorner$ is a wff].
   5. Only expressions obtained by 1–4 are wffs.

III. *Definition of 'Sentence (of LC)'*
   Φ is a sentence iff Φ is a wff with no free variables.

IV. *Recursive Definition of Satisfaction (of a Wff of LC by a Sequence of Classes of Individuals)*
   Let *g* be a denumerable sequence of classes of individuals, and let g_i (where i is a positive integer) be the i-th element of *g*.
   1. *g* satisfies $\ulcorner \overline{\overline{x}}_i \overline{\overline{\subseteq}} \overline{\overline{x}}_j \urcorner$ iff $g_i \overline{\subseteq} g_j$.
   2. *g* satisfies $\ulcorner (\overline{\overline{\sim}}\ \Phi) \urcorner$ iff $\overline{\sim}$ *g* satisfies Φ.
   3. *g* satisfies $\ulcorner (\Phi\ \overline{\overline{\vee}}\ \Psi) \urcorner$ iff *g* satisfies $\Phi\ \overline{\vee}$ *g* satisfies Ψ.
   4. *g* satisfies $\ulcorner (\overline{\overline{\forall}}\ \overline{\overline{x}}_i\ \Phi) \urcorner$ iff $\overline{\forall} g'$ (*g'* differs from *g* at most in its i-th element, $g_i \overline{\supset} g'$ satisfies Φ).

V. *Definition of 'True Sentence (of LC)'*
   A sentence (of LC) is true iff it is satisfied by every sequence *g*.

In setting out to generalize Tarski's example we have to decide which of its features to treat as part of the method and which as peculiar to the specific application. By 'features of Tarski's example' I mean features on various levels of abstraction, from the very simple and technical to the most complex and philosophical. In the next section I will present four questions left open (or partially open) by Tarski's constraints and example: (1) What are "fixed" and "distinguished" constants of Tarski's theory? (2) Is Tarski's theory a "structuralist" or a "reductionist" theory? (3) Is Tarski's theory a genuine correspondence theory? If it is, what entries in his definition are directly responsible for connecting truth to reality? (4) Does Tarski's theory offer a substantive (informative, non-deflationist) account of truth? If not, is it compatible with such an account?

4. Open questions

1. *What Are the Fixed and Distinguished Constants of Tarski's Method?* It is characteristic of the duality of method and application that certain elements of each application are part of the method and others are not, that the former are common to all (or to a large group of) applications and the latter vary from one application to another. I will call the former "*fixed*", the latter "*variable*" elements. The fixed elements of a given method are often treated more informatively, more determinately, more discriminately than its variable elements. When this is the case, I will say that the fixed elements are "*distinguished*". The distinguished status of an element is often manifested by the assignment to it of a distinct and precise application rule, a rule that distinguishes it from all other elements. In contrast, the non-distinguished elements are often treated "en mass" and many of their differences are deemed irrelevant for the method. Understanding a method is, to a large extent, identifying its fixed (and, if applicable, distinguished) elements and knowing how they each operate.

Tarski's example, by itself, does not enable us to determine what the fixed and distinguished elements of his method are. This is due partly to the singleness of his example (its being the only example),[6] partly to its extreme simplicity. In this example each constant belongs to a different syntactic category and as a result differences in the treatment of constants can be predicated on syntactic differences. But syntactic status, too, cannot be established as the dominant factor, since the example provides no occasion to observe how distinct constants of the same syntactic category are treated. Tarski's restriction of his method to languages of the deductive sciences suggests that the logical constants (which are common to all deductive languages) are the fixed (and possibly distinguished) elements of his theory, and this view is supported by his theory's relevance to logic (see "the logical aim" above). But it is not clear whether the restriction to logical languages is heuristic or a matter of principle, whether it is the logicality of '$\sim$', '$\vee$' and '$\forall$' that is responsible for their purported fixity or some other property they all share (e.g., iterativity), what the boundaries of the logical are (see Tarski, 1936b), and whether and in what way Tarski's treatment of the logical constants is more "distinguishing" (more informative and individual) than his treatment of the non-logical constants.

2. *Two Faces of Recursion: Is Tarski's Theory a Reductionist Or a Structuralist Theory of Truth?* – Technically, the recursive method allows us to specify the extension of infinite predicates (predicates with an infinite extension) in a finite manner, based on the inductive structure of elements in a given domain.[7] Philosophically, there are two natural ways of viewing the recursive definition of a predicate P: (i) *the reductionist view* – the definition reduces the conditions under which P applies to objects in general to the conditions under which P applies to structurally simple (atomic) objects; (ii) *the structuralist view* – the definition provides an account of the role played by structural complexity in the satisfaction of P. The two views can be regarded as two sides of the same coin, but they can also be viewed as leading to two different projects involving the definition of P. In the case of truth, the reductionist project seeks to define the general notion of truth by defining the notion of atomic truth, while the structuralist project seeks to explain the role of structural complexity (or a certain kind of structural complexity) in truth. Here the idea is that different kinds of structural complexity are associated with different truth (satisfaction) conditions, and the task of a theory of truth is to spell out these conditions.

The reductionist construal of Tarski's theory (and the theory of truth in general) may assume a narrower or a broader form. More narrowly, the project is purely technical, namely, reducing an infinite task to a finite task; more broadly, the project is embedded in a more comprehensive reductionist program, for example, that of reducing all notions used in science to purely physicalistic notions. The broader reductionist project of truth proceeds in two steps: (i) reduction of the general notion of truth to the notion of atomic truth, and (ii) reduction of the notion of atomic truth to a notion (or set of notions) of the designated kind. One important difference between the structural and the reductionist projects is in the roles played by the atomic and structural entries (the base and recursive entries): In the reductionist project the main philosophical task (that of replacing semantic notions by non-semantic notions of the designated type) is carried out by the atomic entries while the recursive entries play a merely technical role; in the structural project the tables are turned: the main philosophical task is carried out by the structural (recursive) entries, while the atomic entries play an auxiliary role (i.e., that of providing basic arguments for the structural operators).

Neither Tarski's general comments nor his specific example tell us how to view his method. Tarski's example is too elementary to adjudicate between the two approaches, and his general requirement that truth be defined in non-semantic terms could be motivated either by a reductionist goal, as suggested in his 1936a paper, or by general considerations of clarity, consistency and non-circularity, as suggested in his 1933 paper. Had the 'distinguished-nondistinguished' dichotomy coincided with the 'structural-atomic' (or 'atomic-structural') dichotomy, the matter would have been decided, but the former dichotomy was, as we have indicated above, left undecided.[8] In the philosophical literature a reductionist generalization of Tarski's theory is offered by Field (1972), a structuralist generalization by the early Davidson (1965, 1967a, and elsewhere).[9] It is possible, of course, to generalize Tarski's method in a way that is neither reductionist nor structuralist; the generalization I offer in Sher (1998–99) falls under this category.

3. *Is Tarski's Theory a Genuine Correspondence Theory? What Are the Channels of Correspondence in Tarski's Theory?* – Tarski, as we have seen above, conceived of his theory of truth as a classical correspondence theory. This conception he intended to capture by his material adequacy condition, Criterion T. Tarski's formulation of Criterion T, however, was done in linguistic terms: the criterion talks about sentences, names and translations of sentences, not about objects in the world. The question thus arises whether Tarski's notion of truth is not, *in effect*, essentially a linguistic notion.

There are three reasons for taking Tarski's notion of truth to be a correspondence notion: (i) Tarski intended it to be one.[10] (ii) There is no conflict between the claim that Criterion T is formulated in linguistic terms and the claim that it is a *bona fide* correspondence criterion: any objectual statement can be expressed as a statement about language, and the fact that a statement is so expressed does not render it non-objectual. Thus, Tarski says:

> A characteristic feature of the semantical concepts is that they give expression to certain relations between the expressions of language and the objects about which these expressions speak, or that by means of such relations they characterize certain classes of expressions or other objects. We could *also* say (making use of the *suppositio materialis*) that these concepts serve to set up the correlation between the names of expressions and the expressions themselves. (Tarski, 1933, p. 252, first italics mine)[11]

Moreover, it is clear why Tarski chose to formulate his material adequacy condition in linguistic rather than objectual terms: Tarski's strategy of dealing with the correspondence condition as a formal condition expressed in terms of the syntactic notion of derivability means that it is more convenient for him to deal with this condition on the linguistic (syntactic) level than on the objectual level. But this choice of mode of expression does not affect the content of the Criterion. And what the Criterion says, in objectual terms, is that an adequate definition of truth equates the truth of a sentence with the fulfillment of an objectual condition. For example, the truth of the sentence 'Some class of individuals is included in another' is equated with the fulfillment of the corresponding objectual condition, namely, that some class of individuals is included in another. (iii) Finally, the correspondence nature of Tarski's notion of truth is demonstrated by his choice of *satisfaction* as the central semantic notion of his method.[12] Satisfaction is a relation between *objects* (sequences of objects) and linguistic entities (formulas of a given language), hence inherently a correspondence relation.

But how are true sentences connected to reality in Tarski's theory? What entries in the definition of satisfaction are responsible for this connection? It is commonly taken for granted that the atomic entries are the correspondence entries while the structural entries are linguistic: the satisfaction condition for '$\subseteq$' is objectual, but the satisfaction conditions for '$\sim$', '$\vee$' and '$\forall$' are essentially linguistic. This claim, however, has no immediate basis in Tarski's theory. Compare the entries for '$\subseteq$' and '$\forall$' in Tarski's example. The satisfaction condition of '$\subseteq$' is given in terms of *objects* (g_i and g_j) and a *translation* of '$\subseteq$' ; similarly, the satisfaction condition of '$\forall$' is formulated in terms of objects (g and g_i) and the meta-linguistic translation of '$\forall$'. Both satisfaction conditions involve language and the world; why say the one is a correspondence condition, the other not? – The ground for this claim is the belief that '$\overline{\subseteq}$' is an objectual term, while '$\overline{\sim}$', '$\overline{\vee}$' and '$\overline{\forall}$' are linguistic (syncategorematic) terms. But nothing either in Tarski's example or in his general comments on method commits his theory to this view. In (1933) Tarski says nothing about the nature of the logical constants, and the question of whether these constants, and the entries assigned to them in Tarski's definition, are linguistic or objectual, is left open. (Among the well-known analyses of logical constants as objectual operators are

Boole (1854), Frege (1884), Mostowski (1957), Tarski (1966), Lindström (1966), Henkin-Monk-Tarski (1971, 1981), and many contemporary works in model-theoretic semantics.)[13] But even the traditional approach to the "route of correspondence" leaves many options open. Some of these options will be discussed below.

4. *Does Tarski's Theory Offer a Substantive (Informative, Non-Deflationist) Account of Truth? Is It Compatible With Such an Account?* This question will be dealt with in great detail later on. Here I would just like to make a few brief remarks. The question whether a given theory is substantive can be understood in several ways: as a question of whether the given theory satisfies some predetermined standard of substantiveness, whether it fulfills some predetermined task or tasks, whether it says something that for some predetermined group (or number, or percentage) of readers is new, informative, instructive, surprising, explanatory, has interesting ramifications, etc. Tarski, as I have indicated above, did not intend to develop a philosophical theory of the "nature" of truth, and his theory, as reflected in his general comments and example, is intuitively trivial in at least one respect, namely: its T-sentences are all intuitively trivial. Thus, consider the LC sentence '$(\forall x_1)x_1 \subseteq x_1$'. A person who does not know under what conditions '$(\forall x_1)x_1 \subseteq x_1$' is true will not be enlightened by Tarski's answer: '$\ulcorner(\overline{\forall}\overline{x}_1)\overline{x}_1 \overline{\subseteq} \overline{x}_1\urcorner$ is true iff $(\overline{\forall}\overline{x}_1)\overline{x}_1 \overline{\subseteq} \overline{x}_1$'. Nevertheless the possibility of generalizing Tarski's example to a substantive theory in one sense or another is left open. I will discuss one attempt at a substantive generalization of Tarski's theory in Section 7 below.[14]

5. Field's first generalization of Tarski's theory

In his 1972 paper Field offered two reductionist generalizations of Tarski's theory, the one said to capture Tarski's original theory, the other Field's conception of an adequate Tarskian theory.[15] The two theories agree on some of the questions posed above but differ on others: (1) Both theories treat the logical constants as fixed (and weakly distinguished) constants. (2) Both theories are reductionist and in both the intended reduction is physicalistic. (3) Both theories are correspondence theories; but while the two agree that the connection between language and reality takes place in the atomic realm, their account of this connection differs radically. Finally, (4) the two theories diverge on the substantiveness issue: "Tarski"'s theory is a paradigm of a non-substantive reductionist theory, Field's theory (if successful) – a paradigm of a substantive theory.

Field follows Tarski in presenting his construal of his (Tarski's) theory through a single example, which I will refer to as 'T_T'[16]. T_T is more general than Tarski's example in the following respects:

(i) Tarski's example is specific, Field's example is partially schematic. In particular, Field treats non-logical constants schematically, while Tarski treats all constants (both logical and non-logical) non-schematically.
(ii) Tarski's syntax contains only one non-logical constant; Field's syntax contains an open-ended number of non-logical constants.
(iii) Tarski's syntax contains non-logical constants of one category only; Field's syntax contains non-logical constants of three syntactic categories.
(iv) In Tarski's syntax the categories of logical constants and iterative constants coincide; in Field's syntax they do not: Field's syntax contains both logical and non-logical iterative constants.
(v) Tarski's syntax contains only simple singular terms; Field's syntax contains both simple and structurally complex singular terms.

All these differences make for a far greater syntactic generality of Field's example compared with Tarski's. Another difference is the following:

> While Tarski constructed his definition for a language of a particular deductive theory, Field does not mention an object theory. Field, however, assumes each sentence of the object-language (and the corresponding meta-linguistic sentence) has (have) a definite truth value, and this assumption is tantamount to thinking of the object-language as belonging to a particular world-theory.

Below I will use 'T_T' to refer both to Field's general construal of Tarski's theory and to his semi-schematic example of this construal. The intended use of this tag should be clear from the context.[17]

T_T[18]

Object language: L.
Meta-language: ML.

ML has the same structure, principles, and relation to L as in Tarski's example, but the notion of *translation*, or *having the same meaning as*, or *saying the same thing as*, is interpreted (at least with respect to the primitive non-logical constants of L) as *co-extensionality*.

Notation: again I will use '$\overline{\overline{e}}$' to indicate an ML-name of the L-expression e, and '$\overline{e}$' to indicate an ML-expression having the same meaning as e.

I. *Vocabulary of L*
   (a) primitive logical constants: '$\sim$', '$\vee$', '$\forall$'; same meaning as in Tarski's example.
   (b) primitive non-logical constants (represented schematically):
      (i) individual constants – 'c_1', 'c_2', . . . , 'c_k',
      (ii) 1-place 1st-order function symbols – 'f_1', 'f_2', . . . , 'f_m',
      (iii) 1-place 1st-order predicates – 'P_1', 'P_2', . . . , 'P_n',
      where k, m and l are positive integers;
   (c) variables: 'x_1', 'x_2', . . . , ranging over an unspecified domain of objects (treated as individuals).
   (d) punctuation marks: '(',')'.

II. *Inductive Definition of 'Well-Formed Term (Wft, of L)'*
   Let 'i' be a variable ranging over positive integers, and let 't' be a schematic letter representing an arbitrary expression of L.
   1. $\overline{\forall}$i [$\overline{\overline{x}}_i$ is a wft].
   2. $\overline{\forall}$i [i $\leq$ k $\overline{\supset}$ $\overline{\overline{c}}_i$ is a wft].
   3. t is a wft $\overline{\supset}$ $\overline{\forall}$i [i $\leq$ m $\overline{\supset}$ $\ulcorner \overline{\overline{f}}_i(t) \urcorner$ is a wft].
   4. Only expressions obtained by 1–3 are wfts.

III. *Inductive Definition of 'Wff (of L)'*
   Let 'Φ' and 'Ψ' be schematic letters representing arbitrary expressions of L.
   1. t is a wft $\overline{\supset}$ $\overline{\forall}$i [i $\leq$ n $\overline{\supset}$ $\ulcorner \overline{\overline{P}}_i(t) \urcorner$ is a wff].
   2. Φ is a wff $\overline{\supset}$ $\ulcorner (\overline{\overline{\sim}}\ \Phi) \urcorner$ is a wff.
   3. [Φ is a wff $\overline{\&}$ Ψ is a wff] $\overline{\supset}$ $\ulcorner (\Phi\ \overline{\overline{\vee}}\ \Psi) \urcorner$ is a wff.
   4. Φ is a wff $\overline{\supset}$ $\overline{\forall}$i [$\ulcorner (\overline{\overline{\forall}}\ \overline{\overline{x}}_i\ \Phi) \urcorner$ is a wff].
   5. Only expressions obtained by 1–4 are wffs.

IV. *Definition of 'Sentence (of L)'*
   Φ is a sentence iff Φ is a wff with no free variables.

V. *Recursive Definition of Denotation (under* g*)*
   Let A be the intended universe of discourse of L. Let *g* be a denumerable sequence of members of A and let g_i be the i-th member of *g*.
   1. $\overline{\overline{x}}_i$ denotes g_i under *g*.
   2. $\overline{\overline{c}}_i$ denotes $\overline{c}_i$ under *g*.
   3. $\ulcorner \overline{\overline{f}}_i(t) \urcorner$ denotes $\overline{f}_i(\overline{t})$ under *g*, where $\overline{t}$ is the object denoted by t under *g*.

VI. *Recursive Definition of Satisfaction (by* g*)*
   1. *g* satisfies $\ulcorner \overline{\overline{P}}_i(t) \urcorner$ iff $\overline{P}_i(\overline{t})$, where $\overline{t}$ is the object denoted by t under *g*.
   2. *g* satisfies $\ulcorner (\overline{\overline{\sim}}\ \Phi) \urcorner$ iff $\overline{\sim}$ *g* satisfies Φ.
   3. *g* satisfies $\ulcorner (\Phi\ \overline{\overline{\vee}}\ \Psi) \urcorner$ iff *g* satisfies $\Phi\ \overline{\vee}$ *g* satisfies Ψ.
   4. *g* satisfies $\ulcorner (\overline{\overline{\forall}}\ \overline{\overline{x}}_i\ (\Phi) \urcorner$ iff $\overline{\forall}g'$(*g'* differs from *g* at most in g_i $\overline{\supset}$ *g'* satisfies Φ).

VII. *Definition of 'True Sentence (of L)'*
   A sentence (of L) is true iff it is satisfied by every sequence *g*.

Fixed and Distinguished Constants. The clue to identifying the fixed and variable constants of T_T is the distinction between specific and schematic constant symbols: the schematic symbols represent constants which vary from one language to another (within the range of Tarski's method), while the non-schematic symbols represent constants which are fixed across languages. It is clear from Field's example that he identifies fixed constants with logical constants: all the fixed constants of Tarski's method are logical and only logical constants serve as fixed constants. (In Tarski's example the category of logical constants coincides with that of iterative constants, so even if we accept '$\sim$', '$\vee$' and '$\forall$' as the fixed constants of his method, it is impossible to know, based on this example, whether these constants are fixed in virtue of being logical or in virtue of being iterative.) The fixed constants of T_T are, however, distinguished only in a weak sense: each logical constant has a precise, individual satisfaction condition, but this satisfaction condition is intuitively trivial. These conditions essentially say that '**Not** *s*' is true iff **not** (*s* is true), '**Every** x is a P' is true iff **every** object in the universe is a P, etc.

Is T_T committed to a particular choice of logical constants? There is no indication that Field regards Tarski's method as essentially restricted to a particular set of logical constants. It appears to be perfectly compatible with Field's analysis that the set of logical constants include non-standard logical constants, provided, perhaps, that the resulting language would constitute an adequate framework for the systematization of science.

Physicalistic Reduction. Field regards Tarski's project as motivated by the desire to explicate truth in accordance with the methodology of "the unity of science and physicalism" (see Section 2 above). This methodology embeds the project of truth within the broader project of physicalistic reduction, which Field understands as the project of reducing 'chemical facts, biological facts, psychological facts and semantical facts [to distinctly] physical facts' [Field, 1972, p. 91], i.e., as physicalism in the strict sense of the word.[19] Tarski's reduction, on Field's construal, proceeds in three steps:[20]

1. Reduction of the semantic notion of *truth* to the semantic notion of *satisfaction*.
2. Reduction of the semantic notion of satisfaction to the semantic notion of *primitive reference* (the reference of the primitive non-logical constants of the language), which may be broken into three notions: *denotation* (of an object by a primitive name), *application* (of a primitive predicate to an object), and *fulfillment* (of a primitive function by a pair of objects).
3. Reduction of the semantic notions of primitive reference to non-semantic, *physicalistically acceptable notions* (i.e., either logico-mathematical notions or physical notions).

From the point of view of the physicalistic reduction the most significant stage is the last, since it is at this stage that the passage from a physicalistically unacceptable discourse to a physicalistically acceptable discourse takes place.

Correspondence. it is also at this stage that the objectual nature of truth, i.e., its dependence on the way things are in the world, is established. This dependence is based on a correspondence relation between the primitive constants of L and objects (structures of objects) in the world, and this relation is established by means of *lists of (primitive) reference*. These lists correlate each primitive name, 'c_i', of L with an object (its denotation) in the universe of discourse, A, of L; each primitive function, 'f_i', of L with a set of pairs of objects (the set of pairs of objects fulfilling it) in A, and each primitive predicate, 'P_i', of L with a set of objects (the set of objects it applies to, or its extension) in A. Tarski, according to Field, introduces these lists indirectly, namely, via the procedure for translating L into ML. This translation procedure requires that each primitive constant, e, of L, be canonically translated into a constant, $\overline{e}$, of ML, such that (i) $\overline{e}$ contains no semantic terms, and (ii), e and $\overline{e}$ are co-referential.

Trivial Reduction. Tarski's method, according to Field, trivializes the idea of a physicalistic definition of truth. Field explains the sense in which Tarski's definitions are non-substantive using an example from chemistry: Consider the chemical notion of *valence*.

> The valence of a chemical element is an integer that is associated with that element, which represents the sort of chemical combinations that the element will enter into.
>
> [W]e often apply the term 'valence' not only to elements but also to configurations of elements. . . . [I]f we abstract from certain physical limitations on the size of possible configurations of elements . . . , there is an infinite number of entities to which the term 'valence' is applied. But it is an important fact about valence that the valence of a configuration of elements is determined from the valences of the elements that make it up, and from the way they're put together. Because of this, we might try to give a recursive characterization of valence. First of all, we would try to characterize all the different *structures* that configurations of elements can have . . . . We would then try to find rules that would enable us to determine what the valence of a complicated configuration would be, given the valences of certain less complicated configurations that make it up and the way they're put together. If we had enough such rules, we could determine the valence of a given configuration given only its structure and the valences of the elements that make it up. . . . Thus our 'valence definition' . . . would characterize the valence of the complex *in terms of the valences of the simple.*
>
> It would now be possible to eliminate the term 'valence' from [the definition] in either of two ways. One way would be to employ a genuine reduction of the notion of valence for elements to the structural [i.e., physical] properties of atoms. The other way would be to employ the [method of definition by list:]
>
> (∀E) (∀n) (E has valence n ≡ E is potassium and n is +1, or . . . , or E is sulphur and n is –2)
>
> where in the blanks go a list of similar clauses, one for each element. [Ibid., pp. 95–97][21]

The first way represents a substantive reduction of 'valence' to physicalistically acceptable notions, the second – a non-substantive, trivial reduction. Tarski's reduction of truth to physicalistic notions is, according to Field, similar to the second reduction of valence to such notions. The problem lies at the lowest echelon of the recursive procedure, namely, the specification of primitive reference. Here Tarski's *extensional* treatment of the primitive constants is tantamount to definition by lists:

> *Primitive Names*
> $(\overline{\forall}c)\ (\overline{\forall}a)$ [c denotes a $\overline{\equiv}$ [$(c = \overline{\overline{c}}_1\ \overline{\&}\ a = \overline{c}_1)\ \overline{\vee}\ \ldots\ \overline{\vee}$ $(c = \overline{\overline{c}}_k\ \overline{\&}\ a = \overline{c}_k)$]].

Primitive Predicates
$(\overline{\forall}P)\,(\overline{\forall}a)$ [P applies to a $\overline{\equiv}$ [$(P = \overline{\overline{P}}_1\ \overline{\&}\ \overline{P}_1a)\ \overline{\vee}\ \ldots\ \overline{\vee}\ (P = \overline{\overline{P}}_n\ \overline{\&}\ \overline{P}_na)$]],

and similarly for primitive functions. [Based on op. cit., p. 102]

It is these lists that allow us to make the last step from semantic to non-semantic conditions. Thus, if L is a language with 'Boston', 'City' and 'State Capital' as primitive notions, the passage from the semantic condition

g satisfies 'Boston is a city' & g satisfies 'Boston is a state capital'

to the non-semantic condition

Boston is a city & Boston is a state capital

is based on the lists

Den: ⟨'Boston', Boston⟩, . . .
Apl: ⟨'is a city', {Boston, . . .}⟩, ⟨'is a state capital', {Boston, . . .}⟩.

But the trivial nature of these lists trivializes the entire reduction. I will say more about this claim in the next section.

6. Two problems with Tarski's method

Field points out a number of problems with Tarski's method. The two central problems are the relativity of truth to language and the trivialization of truth.

Relativity of Truth to Language. On Field's construal, Tarski's theory is intended to provide a method for specifying the truth conditions of any language within its range in physicalistically acceptable terms. The crucial steps – those in which semantic conditions are replaced by non-semantic conditions – are those dealing with the primitive non-logical constants of a given language, and they essentially involve lists of reference for these constants. But the primitive non-logical constants of Tarskian languages are variable rather than fixed; hence the lists of reference of Tarski's method vary from language to language. This renders the Tarskian notion of truth relative to language: Truth is relative to reference lists, and reference lists are relative to language. In Field's words: '[A Tarskian] truth definition works for a single language only, and so if it 'explains the meaning of' the word 'true' as applied to that language, then for *any* two languages L_1 and L_2, the word 'true' means something different when applied to utterances of L_1 than it means when applied to utterances of L_2!' (Ibid., p. 91). The idea of truth, however, is essentially the same for all languages; hence Tarski's theory fails to capture the intended idea.[22]

Trivialization of Truth. From the point of view of the physicalistic project, the triviality problem is not just a problem of not providing as informative or interesting a reduction as one might wish. The problem is that due to its triviality Tarski's method fails to produce a genuine reduction at all. Thus take an area of discourse that does not yield itself to physicalistic reduction, say *witchcraft discourse*. Provided that the number of primitive constants of this discourse as well as the number of witches, witch spells, etc. is finite, it is possible to generate an *illusion* of a physicalistic reduction of witchcraft discourse by using the list method: 'x is a witch' would be reduced to 'x = Mary ∨ . . . ∨ x = Jean', where Mary, . . . , Jean are all the alleged witches, 'x cast a spell on y' would be reduced to '(x = Mary & y = John) ∨ . . . ∨ (x = Jean & y = Roger)', where ⟨Mary, John⟩, . . . , ⟨Jean, Roger⟩ are all the pairs of alleged witches and their victims, etc. But obviously this would not constitute a genuine physicalistic reduction of witchcraft discourse. (See op. cit., p. 101). In a similar way Tarski's reduction of truth to lists of primitive reference is not a genuine reduction: *co-extensionality* 'is not a sufficient standard of reduction', Field rightly claims. (Ibid., p. 95)

The problem of triviality, however, is not restricted to the physicalistic interpretation of Tarski's theory. It is hard to think of any serious philosophical purpose that would be achieved by a list-based definition of truth: 'it seems pretty clear', Field says, 'that denotation definitions [like those of Tarski] have no philosophical interest whatever'. (Ibid., p. 102)[23]

7. Field's second generalization

Field's solution to the two methodological problems facing Tarski's theory consists in a new treatment of primitive reference. Instead of establishing the reference of the primitive constants of each Tarskian object-language L by means of lists of reference specific to L,

Field proposes that we include in the theory of truth a *general theory of primitive reference*, PR, applicable to all Tarskian object-languages L. This theory will specify the general conditions under which a name denotes an object, a predicate applies to an object, and a function is fulfilled by a pair of objects, in a substantive manner and in accordance with the requirements of physicalism. The result will be a new, revisionist version, or generalization, of Tarski's theory, which I will refer to as *Field's theory of truth*. The revised version of T_T representing Field's theory of truth I will refer to as T_F'.[24] T_F differs from T_T in two respects:

1. The meta-theory of T_F includes a general theory of primitive reference, PR. In particular, T_F includes three general conditions: *D* – a general condition of primitive denotation, *F* – a general condition of primitive fulfillment, and *Ap* – a general condition of primitive application.
2. T_T's primitive-reference schemas,
   (a) $\overline{\overline{c}}$ denotes $\overline{c}$ under *g*,
   (b) $\ulcorner \overline{\overline{f}}(t) \urcorner$ denotes $\overline{f}(\overline{t})$ under *g*, and
   (c) *g* satisfies $\ulcorner \overline{\overline{P}}(t) \urcorner$ iff $\overline{P}(\overline{t})$],
   are replaced by the T_F schemas
   (a*) $\overline{\overline{c}}$ denotes the object possessing the characteristic D_c under *g*,
   (b*) $\ulcorner \overline{\overline{f}}(t) \urcorner$ denotes the object possessing the characteristic $F_{f\bar{t}}$ under *g*, and
   (c*) *g* satisfies $\ulcorner \overline{\overline{P}}(t) \urcorner$ iff $\overline{t}$ possess the characteristic Ap_P,
   where
   – D_c is the physical characteristic determined by the application of *D* to the individual constant c of L;
   – $F_{f\bar{t}}$ is the physical characteristic determined by applying F_f to $\overline{t}$, where F_f is the physical condition determined by the application of *F* to the primitive function f;
   – Ap_P is the physical characteristic determined by the application of *Ap* to the primitive predicate P of L.[25]

Field's proposal, if adequately realized, will offer a full solution both to the relativity problem and to the triviality problem.

The Relativity Problem. PR is a *general* theory of primitive reference and as such it is indifferent to lexical differences between languages. The general conditions (criteria) of reference, *D*, *F*, and *Ap*, are fixed across languages and only their detailed applications vary from language to language. We can compare PR to a general theory of arithmetic operations: the rule of addition does not vary from one pair of natural numbers to another, only its applications do; similarly, the rule of denotation does not vary from one primitive name to another, only its implementation does. Going back to the division of Tarski's method into fixed and variable parts, we may say that on Field's proposed generalization, both the specific principles underlying the semantics of the logical constants and the general principles underlying the semantics of the non-logical constants belong to the fixed, common core of Tarski's method. Since the notion of truth (in this revised version) is reduced to these principles, truth is not relativized to language. The principles of truth are universal (within the range of Tarskian languages); their instantiations alone are particular.

The Triviality Problem. Assuming Field develops a substantive account of primitive reference, the triviality problem will be fully resolved. Using the recursive machinery developed by Tarski, Field will reduce truth to primitive reference, and using his substantive theory of primitive reference, primitive reference will be reduced to physicalistic notions. To the extent that the latter reduction will be substantive (genuine, non-trivial, philosophically significant), Field's entire reduction will be substantive.

Field's Theory of Primitive Reference (PR).[26] Field's success in resolving the two problems depends on the existence of a general and substantive theory of primitive reference. Field considers two accounts of primitive reference: (1) a Russellian "descriptive" account, and (2) a Kripkean "causal" account. The descriptive account was spelled out in detail in Russell (1905, 1918) and elsewhere.[27] The causal account was sketched in Kripke (1972) and elsewhere as part of a negative critique of the Russellian account.[28] Field provides the following outline of the two approaches:

The descriptive theory of reference

> [A] name like 'Cicero' is 'analytically linked' to a certain description (such as 'the denouncer of Catiline'); so to explain how the name 'Cicero' denotes what it does you merely have to explain
>
> (i) the process by which it is linked to the description (presumably you bring in facts about how it was learned by its

> user, or facts about what is going on in the user's brain at the time of the using)
>
> and (ii) how the description refers to what it does.
>
> [To avoid circularity] Russell . . . assumed that the primitives of the language were to be partially ordered by a relation of 'basicness', and that each name except a most basic ('logically proper') name was to be analytically linked to a formula containing only primitives more basic than it. The most basic primitives were to be linked to the world without the intervention of other words, by the relation of acquaintance. (Ibid., pp. 98–99)

The causal theory of reference

> According to [causal] theories [of reference], the fact that 'Cicero' denotes Cicero and that 'muon' applies to muons are to be explained in terms of certain kinds of causal networks between Cicero (muons) and our uses of 'Cicero' ('muon'): causal connections both of a social sort (the passing of the the word 'muons' to laymen from physicists) and of other sorts (the evidential causal connections that gave the original users of the name 'access' to Cicero and give physicists 'access' to muons.) (Ibid., p. 99)

Field rejects the descriptive theory of reference based on criticisms due to Kripke (1971 and 1972) and others. Offering a brief statement of his opposition, Field says:

> This classical view of how names (and other primitives) latch onto their denotations is extremely implausible in many ways (e.g., it says you can refer only to things that are definable from 'logically proper' primitives; it requires that there be certain statements, such as 'If Cicero existed then Cicero denounced Catiline', which are analytic in the sense that they are guaranteed by linguistic rules and are immune to revision by future discoveries). (Ibid.)

Although Field is fully aware of the incomplete nature of Kripke's account and the difficulties involved in completing it, he expresses a guarded optimism with regard to the causal theory:

> The diagnosis that any attempt to explain the relation between words and the things they are about must inevitably lead to either a wildly implausible theory (like Russell's) or a trivial theory (like Tarski's) . . . has become less plausible in recent years through the development of *causal* theories of denotation by Saul Kripke and others . . . , [Kripke] has suggested a kind of factor involved in denotation that gives new hope to the idea of explaining the connection between language and the things it is about. (Ibid., pp. 99–100)

Field's solution is thus contingent upon the development of a general, substantive theory of primitive reference based on the principles of physicalism. In spite of the considerable work devoted to the realization of this project (see Evans (1973), Putnam (1975), Stampe (1979), Devitt (1981) and others), Field appears to have all but given up his hope for a substantive correspondence theory of truth. In his (1986) Field makes the terse comment that 'it has proved extraordinarily difficult to develop the details of an adequate correspondence theory'. (Field, 1986, p. 67) In the same paper Field points to "mistakes" in his earlier paper. For example: 'In Field (1972) I made a mistake in underestimating the value of a disquotational truth-predicate'. (Ibid., p. 64). I will not discuss Field's grounds for changing his view here, but in the next section I will identify a methodological problem affecting reductionist theories of truth in general, and reductionist generalizations of Tarski's theory in particular.

8. The disunity problem

The Reductionist Project: Recapitulation and Clarifications. Reductionist theories of truth aim at reducing the totality of truth conditions of sentences of a given language to conditions of a particular kind, K (e.g., physicalistic conditions). In the case of Tarskian languages the reduction proceeds in three steps:

1. Reduction of the satisfaction conditions of logically complex formulas to the satisfaction conditions of logically simple formulas.
2. Reduction of the satisfaction conditions of logically simple formulas to the reference conditions of primitive non-logical constants.
3. Reduction of the reference conditions of primitive non-logical constants to conditions of type K.

The last step involves the development (or use) of a theory of reference, $\mathcal{R}$, which, in its application to an object-language L, assigns K-ish reference conditions to the primitive constants of L. Referring to $\mathcal{R}$ as a theory of reference, however, is misleading for the following reason: The idea underlying the correspondence conception of truth is that the truth of a sentence depends on two things: (a) what the sentence says about the world, and (b) how the world is. Accordingly, an account of the truth conditions of a given sentence must tell us (i) what the sentence says, and (ii) what conditions must hold in the world for the world to be as the sentence says. I.e., $\mathcal{R}$ has to tell us not just what objects our constants refer to, but also what worldly conditions have to be satisfied for a given sentence to be true. In

the case of a reductionist theory of type K, both (i) and (ii) have to given in terms of type K.

How are (i) and (ii) to be accomplished in a reductionist theory like Field's? – Take, for example, an L-sentence like

Tarski est un logicien et Tarski est un mâle

The recursive steps in Field's definition of truth reduce the truth conditions of this logically complex sentence to the truth conditions of two logically simple sentences

Tarski est un logicien
and
Tarski est un mâle.

But how are the truth conditions of these two sentences to be specified? In the (strict) physicalistic program this task is composed of two tasks:

1. The task of telling what objects (what individual and what properties) the primitive terms 'Tarski', 'est un logicien' and 'est un mâle' refer to.
2. The task of telling what the physical identity conditions of Tarski are and what the physical satisfaction (application) conditions of 'est un logicien' and 'est un mâle' are.

We can explain the difference between these two tasks as follows:

1. The first task involves
   (a) Telling that 'Tarski' refers to Tarski (rather than to Frege or to Tolstoy or . . . .);
   (b) Telling that 'est un logicien' refers to the property of being a logician (rather than to the property of being a biologist or being an artist or . . .) and 'est un mâle' refers to the property of being a male (rather than to the property of being a female, or being a mammal, or . . .).

   As far as this task is concerned, both accounts have to be given in terms reducible to purely physicalistic terms, but it is compatible with this task that 'Tarski' is associated with the *biological* object Tarski, 'est un logicien' is associated with the *cultural* (or *behavioristic*) property of being a logician, and 'est un mâle' is associated with the *biological* property of being a male.
2. The second task involves
   (a) identifying Tarski as a *physical* object (rather than as a biological object or a phenomenological object or a psychological object or . . .).
   (b) stating the *physical* (rather than cultural or biological or . . .) conditions under which an object is a logician or is a male.

The literature on the causal theory of reference sometimes restricts its attention to the first task, but it is essential for the successful execution of the (strict) physicalistic project that both tasks be fulfilled. A similar requirement holds for other reductionist theories of truth.

The Disunity of Truth.[29] Field's reductionist methodology requires that the correspondence relation between primitive terms and reality be accounted for by a single substantive principle.[30] The existence, or even the possibility, of such a principle has, however, never been established. The *disunity claim* is the claim that truth (reference) is too diversified to be captured by a single principle. Reductionist theories of truth in the style of Tarski are especially vulnerable to this claim, since their recursive apparatus decreases the diversity of truth (satisfaction, reference) by no more than one or two factors, namely, the logical and/or iterative factor, but the bulk of diversifying factors is unaffected. Physical, biological, psychological, ethical, mathematical, logical, fictional sentences (terms) all fall in the atomic realm, but the truth (reference) conditions of these elements are, *prima facie*, fundamentally different from each other.

The disunity claim can be formulated in a stronger or a weaker manner according to the imputed degree of disunity and the categoricity of the claim. Using the categories of "strict" and "partial" disunity and "strong" and "weak" claim, I will distinguish four claims: (i) the ***strong*** *claim of* ***strict*** *disunity* states (categorically) no two truths (falsities) can be accounted for based on the same principle;[31] (ii) the ***weak*** *claim of* ***strict*** *disunity* says that we have no good reason to believe that any distinct truths can be accounted for based on the same principle; (iii) the ***strong*** *claim of* ***partial*** *disunity* says that not all truths can be accounted for based on the same principle; and (iv) the ***weak*** *claim of* ***partial*** *disunity* says that we have no good reason to believe that all truths can be accounted for by the same principle. Each of these claims can be relativized to atomic sentences (formulas, terms).

The strong claim of strict disunity is succinctly expressed in the following passage by Blackburn (who, it should be noted, does not endorse it):

> [C]ompare 'is true' . . . with a genuine target of philosophical analysis: 'is conscious', or 'has rights', for example. We investigate these by looking for the principles which determine whether something is conscious, or has rights. These principles are intended to govern any such judgements, so that we get a unified class: the class of conscious things, or things that have rights. Each item in such a class is there because it satisfies the same condition, which the analysis has uncovered. Or, if this is slightly idealized, we find only a "family" of related conditions or "criteria" for application of the terms. Still there is then a family relationship between the members of the class. But now contrast 'is true'. We know *individually* what makes this predicate applicable to the judgements or sentences of an understood language. . . . But these reasons are entirely different [in the case of different judgements or sentences]. There is no single account, or even little family of accounts, in virtue of which each deserves the predicate . . . *There are as many different things to do, to decide whether the predicate applies, as there are judgements to make.* So how *can* there be a unified, common account of the "property" which these quite different decision procedures supposedly determine? . . . The idea that there is [a common property of truth] is an illusion. (Blackburn, 1984, pp. 230–231, italics mine)

Wright (1998) attempts to refute the strong claim of strict disunity by an argument whose implicit import is that of a *reductio ad absurdum*: If truth is subject to the strict strong disunity claim, says Wright, then any concept (or many a concept, or any concept of a certain common kind) is subject to it as well; but obviously not all concepts (not even most concepts/most concepts of the designated kind) are subject to this claim; hence truth is not subject to it either. Wright brings two examples of properties whose satisfaction, like the satisfaction of 'truth', has to do with the particular circumstances of their (potential) satisfiers: (i) the property of having fulfilled one's educational potential, (ii) the property of being twice as old as one's eldest child. Fulfilling one's educational potential means different things for different people, and likewise, being twice as old as one's eldest child means different things for different parents. If you accept the strong disunity claim, Wright says,

> you might just as well say that there is no single thing in which being twice as old as one's oldest child consists (being a *doubletenarian*) since for me it would involve being twice as old as Geoffrey, for Prince Charles being twice as old as William and for Blackburn being twice as old as Gwen.

But, obviously, Wright argues, one can give a general account of both these properties:

> [i] To fulfill one's educational potential is for there to be certain levels of academic attainment such that under certain normal educational conditions it is possible for one to meet them, and such that one has met them.
>
> [ii] To be twice as old as one's oldest child is for there to be some individual of whom one is a father or mother and whose actual age is half one's own.

Wright diagnoses the problem outlined in the citation from Blackburn as applying to all predicates of a certain common kind:

> The general pattern . . . is that of properties whose satisfaction consists of an individual's meeting a condition implicitly involving existential quantification over the right field of a relation. . . . In general, to be the bearer of such a property will be to stand in a relation of a certain kind to an appropriate instance or instances of this implicit quantifier, and the identity of that instance or instances may vary depending on the identity and character in other respects of the bearer in question.

But this does not provide a reason for

> regarding it as an error to suppose, or to try to characterise, a general condition which being F [possessing the property in question] involves satisfying. . . . It is in the *nature* of properties of this general character to admit such variation, and it compromises their unity not at all. (Wright, 1998, p. 13)

I think Wright's argument is successful in showing that *if* the disunity claim applies to truth in virtue of its hidden logical structure, it is implausible. But he has not ruled out the possibility that the strong disunity claim holds of truth in virtue of some other feature having to do with its specific content or semantic nature.[32] Nor has he attempted to refute the partial disunity claims.[33] I believe Wright's argument can be extended to a successful argument against the strong disunity claims regardless of the alleged source of disunity, but the partial disunity claims are more difficult to refute. Indeed, the weak partial disunity claim by itself suffices to raise serious doubts about the reductionist project: If we have no good reason to believe that the truth conditions of all atomic sentences are based on the same principle, the reductionist project is undermined.

The challenge to the reductionist project can be expressed as follows: The reductionist project is based on a division of statements (formulas, terms) into two categories according to their *syntactic complexity*. The project consists in the reduction of the truth (satisfaction, reference) conditions of *syntactically complex* elements to the truth (satisfaction, reference) of *syntactically simple* elements. But there is neither a prac-

tical nor a theoretical guarantee that the truth (satisfaction, reference) conditions of *all* syntactically simple elements are sufficiently unified to be accounted for by a single theory. To see the intuitive force of this challenge, consider two atomic sentences from altogether different fields, say, mathematics and biology ('2 is even' and 'Tarski is dead'). Both the route of reference and the conditions that have to hold in the world for a mathematical statement to be true are (*prima facie*) radically different from those required for a biological statement to be true: A chain of reference connecting an utterance to a number is (*prima facie*) essentially different from one connecting an utterance to a person. (For one thing, a person can, but a number presumably cannot, stand in a causal relation to an utterance.) Similarly, the principles underlying the possession of a mathematical property are (*prima facie*) altogether different from those underlying the possession of a biological property.

The disunity challenge is also a challenge to the parallelism of syntax and semantics. The reductionist project treats all syntactically atomic elements on a par, but the syntactic unity of these elements does not guarantee their semantic unity. Wittgenstein's metaphor of handles in a locomotive (1953: #11–12) offers a clear illustration of this point: Just because all handles in a locomotive look alike, it does not follow that they all perform the same function. Similarly, just because all atomic sentences are syntactically alike, it does not follow that they all have the same truth conditions. Different handles may be visually the same yet operate on different principles, and different sentences may be syntactically the same and yet acquire their truth values based on different principles. The underlying assumption of the reductionist theory of truth, namely, that syntactic unity implies semantic unity, is simply unfounded.

9. Conclusion

Tarski's exposition of his theory leaves many philosophical and methodological questions open, and different answers to these questions lead to different generalizations (interpretations, reconstructions) of his theory. In this paper I have raised four questions left open by Tarski's theory: the question of fixed and distinguished terms, the question of the reductionist vs. the structuralist approach, the question of the "route" of correspondence, and the question of a substantive account of truth. I have examined the answers offered to these questions by two generalizations of Tarski's theory, namely, Field's reductionist-physicalistic generalizations, and I have posed a methodological challenge to the reductionist approach, namely the *disunity problem*. The disunity problem raises two issues: (i) Is there a 1–1 correlation between the syntactic principles underlying the construction of sentences in logical frameworks and the semantic principles underlying their truth conditions? (ii) Is there a ground for believing that the truth conditions of all atomic sentences (the satisfaction conditions of all atomic formulas, the reference conditions of all primitive non-logical terms) can be accounted for based on a single substantive principle? Whether other generalizations of Tarski's theory (in particular "structuralist" generalizations) are immune to this problem is an open question. In Sher (1998–99) I have explored a new approach to the theory of truth, motivated by the disunity challenge.

Notes

[1] The view that Tarski's theory is open to a multiplicity of developments was also expressed by Davidson with respect to Convention T (or, as I will refer to it here, Criterion T): 'Convention T does not settle as much as I thought, and more possibilities for interesting theorizing are open than I had realized'. (1984: xv–xvi).

[2] The philosophical nature of Tarski's enterprise was emphasized by Wolenski and Simon (1989) and Wolenski (1993).

[3] Field is fully aware of the close relationship between Tarski's theory and logic but offers no explanation of this relationship. Davidson's discussion of the relation between the theory of truth and logic requires a more lengthy commentary than I can give here. For a few concise and relatively clear statements by Davidson on this issue see his (1967a, p. 33; 1968, pp. 94–95; 1973, p. 71).

[4] My description of Tarski's general constraints and particular example is not presumed to be complete. I assume the reader is familiar with Tarski (1933) as well as with standard meta-logical notions. I use 'Criterion T' instead of 'Convention T' to indicate that this is a Criterion of material adequacy.

[5] '$\ulcorner \overline{\overline{x}}_i \overline{\overline{\subseteq}} \overline{\overline{x}}_j \urcorner$' names the LC-formula '$x_i \subseteq x_j$', '$\ulcorner (\overline{\overline{\sim}} \Phi) \urcorner$' names an arbitrary negation of LC, etc. Throughout the paper I have tried to direct the reader's attention to the relation between object- and meta-language symbols. Since the meta-linguistic material conditional, for example, is (in principle) definable in LC, I have symbolized it using the upper single bar. For the sake of naturalness and simplicity, however, I have not been completely consistent in my notation.

[6] As indicated above, Tarski did extend his example to languages of higher order, but these extensions do not concern the treatment of constants, which is what we are concerned with here.

[7] It should be noted, though, that not every infinite predicate can be defined by recursion. If the domain of objects is not inductive, or if some objects are generated in more than one way, or if the given predicate does not take into account the structural features of objects in the domain, a recursive definition may not be available.

[8] Note that if we admit the first-order relation of identity as a logical constant, the duality of logical and non-logical constants will no longer correspond to that of structural and atomic elements, unless, of course, we modify our notion of structural (atomic) element. The question of what elements are structural is naturally discussed in connection with Davidson's program. (See, for example, LePore, 1986, and references there.)

[9] In Davidson's program a crucial role is played by the formalization process which precedes the construction of the definition of truth for a given discourse. It is at this stage that the decision on what elements of natural language to count as structural and what structures to attribute to them is made. (See, for example, Davidson, 1967b.)

[10] Tarski continued to see his theory of truth as a correspondence theory throughout his career. See Tarski (1936a, 1944, 1969).

[11] The medieval distinction between 'suppositio materialis' and 'suppositio formalis' is opposite to what one would expect today: In the suppositio formalis mode we view a term as standing for the object it "signifies", in the suppositio materialis mode we view it as a linguistic expression.

[12] Tarski, as we have seen above, explained the need for choosing a different notion from 'truth' as the basic notion of his method based on technical considerations. But the choice of 'satisfaction' is clearly motivated by the goal of capturing the classical (correspondence) conception of truth.

[13] For a generalization of Tarski's theory in which the logical entries are construed as correspondence entries see Sher (1998–99). The objectual construal of the logical constants in this generalization is based on Sher (1991, 1996).

[14] The substantive generalization I will discuss below, due to Field (1972), involves intuitively informative T-sentences. Davidson (1984, p. xv), however, rightly points out that there are other ways a Tarskian theory of truth can be non-trivial than having non-trivial T-sentences.

[15] For an excellent analytic description of Field's theories see McDowell (1978).

[16] Field refers to this example as 'T2'.

[17] Due to the greater generality of Field's example and the greater specificity of his discussion (of the issues we are concerned with here), it is quite clear what his general construal of Tarski's theory is intended to be (at least as far as our concerns go).

[18] My presentation of Field's example diverges from his in a few minor technical and notational matters.

[19] It is highly unlikely that Tarski conceived of his method as constrained by an extreme form of physicalism, one which rejects, for example, the autonomy of biological discourse. Tarski's example reduces the truth of class-theoretical statements to the inclusion of classes, and there is nothing specifically physicalistic about this reduction. This reduction is compatible with vitalism (the theory that reality is ultimately biological) and pythagoreanism (the theory that reality is ultimately mathematical) and phenomenalism (the theory that reality is ultimately phenomenal) as much as with physicalism (the theory that reality is ultimately physical), and as such it does not constitute an example of a physicalistic reduction. Had Tarski been interested in a specifically physicalistic account of truth he would have chosen a different example (or would have given different truth conditions – for example, Millian truth conditions – to the sentences in his example). It is therefore more reasonable to assume that Tarski was aiming at an account of truth consistent with the "scientific spirit" in general, rather than with physicalism in the strict sense. (Soames (1984) also argues against Field's physicalistic interpretation of Tarski's theory.) Be that as it may, a strict physicalistic generalization of Tarski's theory is compatible with both Tarski's example and his general comments. Moreover, everything Field says about Tarski's theory is valid with respect to any (reasonable) version of reductionism, including one that interprets physicalism as scientific rigor (compatibility with the scientific spirit), as suggested above.

[20] We may think of steps 2 and 3 as a single step.

[21] I have mixed the order of passages in this citation. I have also left out a few finer points.

[22] In recent decades the relativity to language problem has been treated as one of the major problems with Tarski's theory. See, for example, Davidson (1990).

[23] Field uses 'denotation' where I use 'reference'.

[24] My 'T_F' combines Field's T1 and his projected theory of primitive reference. This difference explains why and how my formulation of the basic schemas differs from his.

[25] Using Field's style of definition, the progression from T_T's primitive-reference schemas to T_F's can be represented as follows (omitting (A) 1):

T2's primitive-reference schemas (Field, 1972, p. 89):

(A) 2 'c_k' denotes$_s$ $\overline{c}_k$

3 $\ulcorner f_k(e) \urcorner$ denotes$_s$ an object a iff

(i) there is an object b that e denotes$_s$, and

(ii) a is $\overline{f}_k(b)$

(B) 1 $\ulcorner p_k(e) \urcorner$ is true$_s$ iff

(i) there is an object a that e denotes$_s$, and

(ii) $\overline{p}_k(a)$.

T1's primitive-reference schemas (Ibid., p. 85)

(A) 2 'c_k' denotes$_s$ what it denotes

3 $\ulcorner f_k(e) \urcorner$ denotes$_s$ an object a iff

(i) there is an object b that e denotes$_s$, and

(ii) 'f_k' is fulfilled by $\langle a, b \rangle$

(B) 1 $\ulcorner p_k(e) \urcorner$ is true$_s$ iff

(i) there is an object a that e denotes$_s$, and

(ii) 'p_k' applies to a.

T_F's primitive-reference schemas (based on the present text)

(A) 2 'c_k' denotes$_s$ an object a iff a has the physical characteristic D_{c_k}

3 $\ulcorner f_k(e) \urcorner$ denotes$_s$ an object a iff

(i) there is an object b that e denotes$_s$, and

(ii) $\langle a, b \rangle$ has the physical characteristic F_{f_k}

(B) 1 $\ulcorner p_k(e) \urcorner$ is true$_s$ iff

(i) there is an object a that e denotes$_s$, and

(ii) a has the physical characteristic Ap_{p_k}.

[26] Field's discussion of primitive reference requires that we conceive of utterances rather than of sentences as bearers of semantic

properties. The difference between the two kinds of bearers is, however, orthogonal to the issues we are discussing; therefore, for the most part I will continue to talk of names and sentences as reference – and truth – bearers. It should be clear from the context when bearers have to be thought of as utterances.

[27] See, for example, Lewis (1946) and, for a more modern version of the descriptive account, Searle (1958).

[28] See, for example, Donnellan (1966, 1970), Kripke (1971), and Putnam (1962a, 1962b, 1970).

[29] An earlier, more general discussion of this problem (referred to as the problem of a substantive theory of truth) appears in Sher (Forthcoming).

[30] One may wish to weaken the requirement by replacing 'a single substantive principle' by 'a unified collection of substantive principles'. However, the challenge I will shortly pose can be posed to the weaker requirement as well.

[31] More precisely, the claim is that the truth conditions of no two sentences differing in content (according to some reasonable criterion) can be accounted for in a substantive manner based on the same principle.

[32] Blackburn may have had such an argument in mind. Note that the properties of being conscious and having rights, which according to his argument can (each) be accounted for by a single principle, may also be interpreted as falling under the category singled out by Wright.

[33] Note that Wright himself (1992, 1998) is moderately pluralist about truth, and his pluralism is naturally understood as motivated by something like the weak version of the partial disunity claim.

References

Blackburn, S.: 1984, *Spreading the Word: Groundings in the Philosophy of Language*, Oxford: Oxford University Press.

Boole, G.: 1854, *An Investigation of the Laws of Thought on which are Founded the Mathematical Theories of Logic and Probabilities*, New York: Dover Publications, 1958.

Davidson, D.: 1965, 'Theories of Meaning and Learnable Languages', Davidson 1984, 3–15.

Davidson, D.: 1967a, 'Truth and Meaning', Davidson, 1984, 17–36.

Davidson, D.: 1967b, 'The Logical Form of Action Sentences', *Essays on Actions and Events*, Oxford: Oxford University Press, 1980.

Davidson, D.: 1968, 'On Saying That', Davidson, 1984, 93–108.

Davidson, D.: 1973, 'In Defence of Convention T', Davidson, 1984, 65–75.

Davidson, D.: 1984, *Truth and Interpretation*, Oxford: Oxford University Press.

Davidson, D.: 1990, 'The Structure and Content of Truth', *The Journal of Philosophy* **87**, 279–328.

Devitt, M.: 1981, *Designation*, New York: Columbia University Press.

Donnellan, K. S.: 1966, 'Reference and Definite Descriptions', *The Philosophical Review* **75**, 281–304.

Donellan, K. S.: 1970, 'Proper Names and Identifying Descriptions', *Syntheses* **21**, 335–358.

Evans, G.: 1973, 'The Causal Theory of Names', *Aristotelian Society Supplementary Volume* **47**, 187–208.

Field, H.: 1972, 'Tarski's Theory of Truth', Platts 1980, 83–110.

Field, H.: 1986, 'The Deflationary Conception of Truth', in G. MacDonald and C. Wright (Eds.), *Fact, Science, and Morality*, Oxford: Basil Blackwell.

Frege, G.: 1884, *The Foundations of Arithmetic*. Tr. J. L. Austin, Evanston: Northwestern University Press, 1968.

Henkin, L., J. D. Monk and A. Tarski: 1971, *Cylindric Algebras*, Part I, North Holland.

Henkin, L., J. D. Monk and A. Tarski: 1981, 'Cylindric Set Algebras and Related Structures', *Cylindric Set Algebras. Lecture Notes in Mathematics* **883**, Berlin: Springer-Verlag, pp. 1–129.

Kripke, S.: 1971, 'Identity and Necessity', in M. K. Munitz (Ed.), *Identity and Individuation*, New York: New York University Press, pp. 135–164.

Kripke, S.: 1972, *Naming and Necessity*, Cambridge: Harvard University Press, 1980.

LePore, E.: 1986, 'Truth in Meaning', in E. LePore (Ed.), *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, Oxford: Basil Blackwell, pp. 3–26.

Lewis, C. I.: 1946, *An Analysis of Knowledge and Valuation*, La Salle: Open Court.

Lindström, P.: 1966, 'First Order Predicate Logic with Generalized Quantifiers', *Theoria* **32**, 186–195.

McDowell, J.: 1978, 'Physicalism and Primitive Denotation: Field on Tarski', Platts 1980, 111–130.

Mostowski, A.: 1957, 'On a Generalization of Quantifiers', *Fundamenta Mathematicae* **44**, 12–36.

Platts, M. (ed.): 1980, *Reference, Truth and Reality: Essays on the Philosophy of Language*, London: Routledge & Kegan Paul.

Putnam, H.: 1962a, 'It Ain't Necessarily So', *The Journal of Philosophy* **59**, 658–671.

Putnam, H.: 1962b, 'Dreaming and "Depth Grammar"', Putnam 1975b, 304–324.

Putnam, H.: 1970, 'Is Semantics Possible?', Putnam 1975b, 139–152.

Putnam, H.: 1975a, 'The Meaning of "Meaning"', Putnam 1975b, 215–271.

Putnam, H.: 1975b, *Mind, Language and Reality*, Cambridge: Cambridge University Press.

Russell, B.: 1905, 'On Denoting', Russell 1956, 39–56.

Russell, B.: 1918, 'The Philosophy of Logical Atomism', Russell 1956, 175–281.

Russell, B.: 1956, *Logic and Knowledge: Essays 1901–1950*. Ed. R. C. Marsh. London: George Allen & Unwin.

Searle, J.: 1958, 'Proper Names', *Mind* **67**, 166–173.

Sher, G.: 1991, *The Bounds of Logic: A Generalized Viewpoint*, Cambridge: MIT Press.

Sher, G.: 1996, 'Semantics and Logic', in S. Lappin (Ed.), *Handbook of Contemporary Semantic Theory*, Oxford: Blackwell Publishers, pp. 509–535.

Sher, G.: 1998–99, 'On the Possibility of a Substantive Theory of Truth', *Synthese* **177**, 133–172.

Stampe, D. 1979, 'Towards a Causal Theory of Linguistic Representation', in P. A. French, T. E. Uehling, Jr. and H. K. Wettstein (Eds.), *Contemporary Perspectives in the Philosophy of Language*, Minneapolis: University of Minnesota Press, pp. 81–102.

Tarski, A.: 1931, 'On Definable Sets of Real Numbers', Tarski 1983, 110–142.

Tarski, A.: 1933, 'The Concept of Truth in Formalized Languages', Tarski 1983, 152–278.
Tarski, A.: 1936a, 'The Establishment of Scientific Semantics', Tarski 1983, 401–408.
Tarski, A.: 1936b, 'On the Concept of Logical Consequence', Tarski 1983, 409–420.
Tarski, A.: 1944, 'The Semantic Conception of Truth and the Foundations of Semantics', *Philosophy and Phenomenological Research* **4**, 341–376.
Tarski, A.: 1954, 'Contributions to the Theory of Models', I and II. *Koninkl. Med. akad. Wetensch. Proc.* Ser. **A 57**, 572–588.
Tarski, A.: 1955, 'Contributions to the Theory of Models', III. *Koninkl. Med. akad. Wetensch. Proc*. Ser. **A 58**, 56–64.
Tarski, A.: 1966, 'What are Logical Notions?', in J. Corcoran (Ed.), *History and Philosophy of Logic* **7** (1986), 143–154.
Tarski, A.: 1969, 'Truth and Proof', *Scientific American* **220**(6), 63–77.
Tarski, A.: 1983, *Logic, Semantics, Metamathematics*, 2nd Edition. Tr. J. H. Woodger. Ed. J. Corcoran. Introduction: J. Corcoran. Indianapolis: Hackett.
Tarski, A. and R. L. Vaught: 1957, 'Arithmetical Extensions of Relational Systems', *Compositio Math* **13** (1958), 81–102.
Vaught, R. L.: 1974, 'Model Theory Before 1945', in L. Henkin et al. (eds.), *Proceedings of the Tarski Symposium*, Providence: the American Mathematical Society, pp. 153–172.
Wittgenstein, L.: 1953, *Philosophical Investigations*. Tr. G. E. M. Anscombe, 3rd Ed., Oxford: Basil Blackwell, 1967.
Wolenski, J.: 1993, 'Tarski as a Philosopher', in F. Coniglione et al. (Eds.), *Polish Scientific Philosophy: The Lvov-Warsaw School*, Amsterdam: Rodopi, pp. 319–338.
Wolenski, J. and P. Simon: 1989, 'De Veritate: Austro-Polish Contributions to the Theory of Truth from Brentano to Tarski', in K. Szaniawski (Ed.), *The Vienna Circle and the Lvov-Warsaw School*, Kluwer Academic Publishers, pp. 391–442.
Wright, C.: 1992, *Truth and Objectivity*, Cambridge: Harvard University Press.
Wright, C.: 1998, 'Truth: A Traditional Debate Reviewed', Manuscript.

University of California
San Diego