Cambridge University Press
978-1-108-42254-3 — The Semantic Conception of Logic
Edited by Gil Sagi , Jack Woods
Excerpt
More Information

The Semantic Conception of Logic: Problems and Prospects

Gil Sagi and Jack Woods

The semantic tradition makes languages and their interpretations the objects of formal study. It has flourished through the development of model theory, initially used by Tarski for the formal explication of the notions of truth and logical consequence. Since the midst of the twentieth century, model theory has had tremendous impact in mathematics, computer science, linguistics, and philosophy. In mathematics, model theory started as a foundational discipline and was typically concerned with the study of consistency, compactness, and completeness. Since then, it has turned its focus to the systematic organization of mathematical theories through applications in various fields from number theory to algebraic geometry.[1]

In computer science, the study of the syntax–semantics interface is paramount in both theoretical and applied topics. The study of formal languages in computer science and of computational models more generally is rooted in the semantic tradition. Database theory and specifically finite model theory are examples of fields using high-powered model theory in computer science.[2] In linguistics, model theory permeates modern-day natural language semantics exemplified in the Montagovian tradition. Philosophy interacts with all the aforementioned disciplines, but also has its own connections to model theory in metaphysics and philosophy of science, in particular in studying reference and its connection to realism.[3]

The semantic tradition in the philosophy of logic, which is our primary concern here, develops and scrutinizes the Tarskian definitions of truth and consequence using philosophical notions as well as mathematical tools, including model theory itself. Since model theory prevails in

[1] For the shifting role of model theory in mathematics, see Manders (1987); Baldwin (2018).
[2] For a survey of some of the interaction between model theory and computer science, see Makowsky (1995).
[3] For an extensive presentation of the uses of model theory in philosophy, see Button and Walsh (2018).

so many other fields, it's important to reflect on its presumptions, how it works, and whether it depends in any significant way on facts from other disciplines, such as science or mathematics. These questions permeate this collection as the reader will shortly see.

The wider historical context of the present collection emerges at the turn of the twentieth century, which marks the advent of formal tools in logic. The first stabs at contemporary rigor in logic involved formalizing syntactic notions, such as formation and transformation rules for formulas: the former concern defining well-formed formulas, and the latter concern laws of deduction. Such rules are completely indifferent to the intended semantic content of their objects. The priority of syntactic approaches over semantic approaches seems to have arisen from the suspicion that semantic notions weren't well understood, were rather metaphysical (when that was a term of disapprobation), and smelt of paradox.

Tarski's "The concept of truth in formalized languages" (1933) changed all of this. It provided a rigorous account of truth for sentences which can then be used to define other important semantic notions like consistency and logical consequence. This went a long way towards demystifying semantic approaches to logical properties. It also paved the way for Tarski's "On the concept of logical consequence" (1936) where he analyzed logical consequence in terms that roughly corresponds to the model-theoretic approach we use today.[4]

This definition works as follows: divide expressions into 'logical' expressions and 'non-logical'. Then a sentence φ follows from a collection of sentences Γ if, no matter how we (uniformly) interpret the meaning of the non-logical expressions contained within φ and Γ, φ is true in a model $\mathfrak{M}$ whenever all of Γ are.[5] So 'there are cats' follows from 'there are cats and dogs' since no matter what 'cats' and 'dogs' mean, the first sentence is true whenever the second is. We don't vary the meaning of the quantifier 'there are' since it's intuitively a logical expression.

[4] There are some interpretational issues surrounding the relationship between the Tarskian conception of consequence and our current conception. These tend to center around whether Tarski already was thinking in terms of "varying domains" or whether his account of consequence drew upon a fixed background of objects. See Ray (1996); Gómez-Torrente (1996, 2009); Mancosu (2006, 2014). One may also note that in 1936, Tarski's definition uses an essentially stronger metalanguage, whereas later Tarski and Vaught (1957) definitions are couched in set theory, and do not resort to an essentially stronger metalanguage.

[5] There's been a lively discussion about exactly what this definition requires. In particular, there's a question about the breadth of the necessity captured by 'no matter how'. See Sher (1991), Hanson (1997), and Gómez-Torrente (1999) for discussion.

This definition presumes that we can divide the logical expressions from the non-logical, but it does not itself give a way to do that. The topic of logicality (of expressions) is thus central in the philosophical literature on the semantic conception of logic. In Tarski's initial work, he expressed skepticism that there was a fixed way of doing so, noting that expressions like the set-theoretic relation of inclusion, were sometimes, but only sometimes, treated as logical. His idea seemed to have been that we have, at best, a definition of logical consequence relative to a chosen set of privileged expressions.

Thirty years later, in "What are logical notions?" (a lecture delivered in 1966 and published in 1986), Tarski addresses the issue of logicality from another perspective. He proposes a criterion of *logical notions*, set-theoretic entities which constitute the distinguished subject matter of logic among other mathematical disciplines, in terms of their stability under transformations of the domain. These notions can model the meanings of expressions in a language, and they are susceptible to precise mathematical characterization. The intuitive model-theoretic meaning of 'there are' – the set of all non-empty subsets of the domain – doesn't change when we permute the domain. Although Tarski does not make the connection himself, it is reasonable to identify logical terms as those denoting logical notions, at least at first pass.[6]

The idea is simple: if the meaning of some predicate or quantifier can be disturbed by a permutation, then it somehow depends on the nature of the objects it's about. Any expression that is so worldly doesn't have the requisite 'formal' character that's supposed to characterize logical expressions (McGee 1996). So this formal property of invariance under permutations is used to test for the intuitive property of formality which characterizes logical expressions. Even if one doesn't think this exhausts the analysis of logicality, it's a very plausible necessary condition on an expression being logical.

This kind of definition of logicality can be made precise by assigning each expression a "meaning" drawn from a type-hierarchy built over a domain of objects. We assign the predicate 'cat' the function from the domain to {T,F} which returns true if an object is a cat and false otherwise. Simplifying, we treat the meaning of 'cat' as the set of cats in the domain. We assign the two place predicate '=' the set of ordered pairs $\langle d, d\rangle$ for d

[6] Corcoran, in his editorial introduction to Tarski (1966/86), presents Tarski as filling the lacuna left in his account of logical consequence of 1936. However, for reservations of this interpretation of Tarski, see Sagi (2017).

in the domain. And so on. Then it's easy to see that '=' is a logical expression since the set of ordered pairs $\langle \pi(d), \pi(d) \rangle$ for d in the domain and π a permutation is just the set of ordered pairs $\langle d, d \rangle$ from the domain. Whereas there's no guarantee that $\pi(d)$ for d a cat is a cat.

The *Tarski-Sher thesis* identifies the logical expressions as those which denote logical notions in the sense just defined – so invariance under certain transformations is both a necessary and sufficient condition for logical notions, and denoting logical notions is a necessary and sufficient condition for something to be a logical expression, if it is adequately incorporated in a syntactic-semantic system of logical consequence.[7] Several writers, among them Vann McGee, William Hanson, Mario Gómez-Torrente, and Timothy McCarthy, have criticized this approach for its "extensionality": invariance criteria look only at denotations, and are blind to the level of sense, or of what connects an expression to its denotation.[8]

There are further concerns about this approach. As Feferman pointed out (1999), the formal approaches to identifying logical expressions rely on the existence (or not) of certain permutations and isomorphisms. Yet 'isomorphism' is not an "absolute" property, so changing the background set-theoretic assumptions can change whether or not a notion is logical. Such problems generalize, showing that the Tarskian analysis of logical consequence is dependent on mathematical facts. Since facts about logical consequence are supposed to be prior to mathematical facts, these "entanglements" have bothered many theorists.[9]

On the positive side, logicians have identified interesting connections between a notion being logical and that notion being definable (van Benthem 1982; McGee 1996), between invariant notions and interpolation theorems (Barwise and van Benthem 1999), and between invariance properties and compositionality (Keenan and Westerstahl 1997).[10] Linguists have also identified interesting connections between an expression being logical, in the sense of being invariant under permutation, and grammatical properties.[11] So regardless of the status of the worries mentioned above, the semantic conception of logic has proved a fruitful way of investigating

7 See Sher's chapter in this volume for her latest take on the issue. Griffiths and Paseau (2022) is a forthcoming book on logical monism and the Tarski-Sher thesis.

8 See Gómez-Torrente's contribution to this volume, as well as his (2002), along with McGee (1996); McCarthy (1981); Hanson (1999); Warmbrōd (1999); Feferman (1999).

9 See Florio and Incurvati's contribution to this volume for an investigation of similar "entanglement" phenomena.

10 See van Benthem's contribution to this volume for a survey of other interesting connections.

11 See Chierchia's contribution to this volume, in particular, for an overview of these connections.

a number of issues. Moreover, the semantic conception itself might be separable from the approach just discussed to logicality (see Gómez-Torrente's contribution to this volume.)

In the present volume, we see the semantic account applied to modal notions (van Benthem), fragments of pure mathematics (Florio and Incurvati), grammaticality (Abrusán et al.; Chierchia), validity in natural language (Glanzberg), and purely formal logical validity (Paseau and Griffiths; Cook). We see the broader implications of the semantic conception explored in the contributions of Bonnay and Speitel, Gómez-Torrente, Sher, Madison Mount, and Zinke. Together these provide a balanced and nuanced picture of both the applications and limitations of the semantic approach.

Contributions

Our volume opens on the use of invariance criteria to characterize logicality. Gila Sher's essay "Invariance and Logicality in Perspective" introduces her seminal approach to using invariance properties to characterize formality and thereby logicality (1991). Sher's view is that logical constants are expressions that pick out properties which are isomorphism invariant as well as satisfying a few other conditions concerning how they are picked out. This approach, as mentioned above, has been very influential over the last 30 years. Sher then surveys how developments of the last 35 years have affected her project. This involves distinguishing her approach from pragmatic alternatives (represented by Hanson (1997) and Gómez-Torrente (2002, this volume)) and argues that her approach maintains desirable connections between logicality, formality, and necessity.

Sher's essay is followed immediately by Gómez-Torrente's defense of his pragmatic alternative. Gómez-Torrente argues that the semantic conception of logic, as described above, works perfectly well even with an account of logicality shot through with local pragmatic choices. In developing his pragmatic alternative, he usefully revisits some of his earlier objections to Sher's non-pragmatic account of logicality (2002), such as worries about properties which are invariant, but which are necessarily denoted by expressions which do not seem logical. Consider 'is a regular heptahedron' or 'is a male widow'. He then suggests that there are a number of other marks of logicality, such as broad applicability in our systematic theorizing. His pragmatic suggestion leaves the notion of logicality intentionally a bit vague, rendering as clearly logical those expressions which score sufficiently well on all the marks, and clearly non-logical those which score sufficiently badly on them. This allows a more degreed notion of the logicality of expressions.

We then turn to the contribution of Bonnay and Speitel which attempts to solve some of the problems with the definition of logicality mooted above by Sher and criticized by Gómez-Torrente. While Bonnay and Speitel agree with treating invariance under bijections as a necessary condition on logicality, they hold that it needs to be augmented with certain syntactic constraints. Drawing on an idea dating back to Carnap (1943), they want logical expressions to be *categorical* in the sense of their meaning being uniquely given by their inferential role.[12] Their overall suggestion is that logical expressions are those which denote a unique bijection-invariant notion by means of their inferential role.[13] This novel suggestion seems a worthy contender for a mixed semantic-syntactic account of logicality, improving (as they note) on another suggestion for a mixed account by Feferman (2015).

Most of the discussion of logicality concerns *extensional* contexts, like those found in mathematics and other informal applications of logic. Yet if we truly want an account of logicality that applies to (rigorized) natural language, we should be willing to consider intensional contexts as well (as argued in Woods 2016). Madison Mount's contribution takes up this challenge, developing an account of logicality in the context of the intensional type theory of Muskens (2007). The account is a useful first step towards developing an intensional criterion of logicality, especially as Madison Mount gives a careful treatment of the challenges and technology necessary to adequately extend invariance-style criteria from extensional to intensional contexts.

Part I closes with Cook's contribution. Cook is concerned with the so-called "validity paradox". This problem, which has gained a lot of contemporary attention, concerns what happens when you add an expression '$\Rightarrow$' expressing logical consequence to our language. Under a number of natural rules governing this expression, and the assumption that these rules are "logically valid", the resulting logic is inconsistent (as Cook (2014) and Ketland (2012) demonstrated in prior work). Here, Cook additionally shows that the meaning of '$\Rightarrow$' is permutation invariant only if it's trivial (in the sense that if anything $\Rightarrow$ anything, then everything $\Rightarrow$ everything). These two observations motivate Cook to develop a more permissive notion of logicality which maintains much of the interest of

12 This meaning of 'categorical', though related to the more usual notion, is a bit narrower.
13 Note that this suggestion allows that other notions would satisfy the inferential role of these expressions, but that these notions must not be bijection-invariant.

the permutation invariance account while allowing '⇒' to be a consistent and non-trivial logical expression. He closes his contribution by mulling on whether the resulting notion of logicality is too permissive – and what this might mean for those pushing the validity paradox as a significant problem for ordinary approaches to logical consequence.

The volume then turns to further basic aspects of the semantic approach. Van Benthem's chapter extends the philosophical discussion of invariance to a broader look at the functioning of semantics in logic. In particular, he discusses connections between the various notions of invariance used in logic and definability in different formal languages, and the fundamental further issue of coherence: how the resulting different views of structure can still be connected. A second main theme is the ubiquitous entanglement of invariance with inference, where, for instance, model-theoretic preservation and interpolation theorems describe a generalized notion of valid inference, going from truths in one model to truths in other suitably related models. Finally, the chapter connects to other fields by drawing attention to the pervasive role of computation and games beneath the surface of semantics.

Incurvati and Florio investigate a problem mentioned above, the entanglement of logic with mathematics. It's a well-known result (see Shapiro 1991) that there's a sentence of full second-order logic (with a standard model-theoretic semantics[14]) that's valid just in case Zermelo-Fraenkel set theory with the axiom of choice (ZFC) proves the continuum hypothesis (CH). And there is another sentence of full second-order logic that's valid just in case ZFC proves the negation of CH. Yet both CH and its negation are provably consistent with ZFC. So the validities of second-order logic seem to determine what ZFC proves, violating what Incurvati and Florio call "neutrality principles". These principles, which claim logic should be dialectically and informationally neutral, are very intuitive, so these *overgeneration* arguments are rather troubling. Incurvati and Florio defend full second-order logic by arguing that the proponent of the semantic conception of logic can adopt a *higher-order* model theory instead of a first-order one. Once they've done that, the troubling overgeneration arguments disappear. They close by exploring whether related arguments can be made against the higher-order approach, tentatively concluding that it doesn't appear to be so.

14 This means that bound function and relation variables range over all the functions and relations on the domain, not just some collection of such. The latter sort of semantics is typically called a "Henkin semantics".

Paseau and Griffiths's contribution explores an issue we've neglected so far. It's natural to take logical truth to be truth in virtue of logical form. The Tarskian approach to logical truth – truth in virtue of logical form – analyzes logical form in terms of holding fixed the logical expressions of a language. It uses the logical expressions occurring in a language to explicate logical form, as explained above. But the distinction between logical and non-logical expressions already presumes another division, a principled division of expressions of the language into grammatical categories. Paseau and Griffiths investigate this presumption in the context of specifying the logic of logical truth, understood as truth in virtue of logical form. That is, in the context of investigating which axioms hold of the concept of logical truth. They show that this depends on whether we take a very fine-grained or coarse-grained approach to treating two expressions as being of the same rough grammatical category. For instance, if G and F are atomic (or basic) predicates, then when we have a claim like 'it's logically true that "$Fa \vee \neg Fa$"', we need to consider whether we mean this to entail that 'it's logically true that "$(Fa \wedge Ga) \vee \neg(Fa \wedge Ga)$"' or only substitutions of atomic predicates like G for F. As they show, the logics of logical truth that result from each way of graining are quite different, one being the normal modal logic $S5$, the other perhaps $S4$. Furthermore, if 'it's logically true that' is not a logical constant, the resulting logic is non-normal. Their contribution helps to show exactly what's being presumed in the background of even something as seemingly straightforward as the semantic approach.

Zinke rounds off this Part 2 of the volume with her related discussion of neglected presuppositions of the semantic approach. As she reminds us, the Tarskian approach mooted above makes substantial assumptions about what sort of reinterpretations of the non-logical terms are admissible. These assumptions give rise to presumptions about the division between analytic and logical truths as well as presumptions about the division between metaphysical and logical truths. Semantic entanglement of this kind is problematic without a rationale since the Tarskian account is supposed to give something like the most basic *formal* truths that make no presumptions about meaning or the world. Zinke explores a number of potential rationales for these presumptions and finds them all wanting. This leaves the proponent of the semantic approach with a challenge: find a rationale for what is presumed or accept that the division between logical and non-logical expressions, and thereby the division between logical and non-logical truths and implications, is somewhat arbitrary. Albeit in a way that usually goes unnoticed.

The third and final Part of our volume addresses connections between the semantic approach to logic and natural language. Glanzberg opens Part 3 by discussing the relationship between logic and the semantics of natural language. His particular focus is on formal model-theoretic semantics for natural language. His earlier work (Glanzberg 2015) attacks the idea that there's a neat connection between logic and the semantics of natural language by arguing that natural language "entailments" either aren't formal or they aren't necessary. His essay here is more constructive, describing ways in which model-theoretic semantics *can*, in fact, be seen as modeling aspects of natural language. The most novel of these is the idea that model-theoretic semantics can function like analogical models in the sciences. That is, as formally precise structures having certain structural properties that are analogous to some kind of structure in the modeled phenomenon. The match need not be perfect, but in order to be successful something *like* the structural property must be present in the modeling target. This is a development of the attractive "logic as modeling" approach first mooted by Shapiro (1998) and developed in Cook (2002).

Chierchia introduces and defends the idea that many cases of ungrammatical sentences can be explained in terms of their logical falsity (an idea originally suggested and developed by Gajewski (2002)). Drawing on recent work by Del Pinal (2019), Chierchia shows how we can identify a class of sentence – the *G-trivial* sentences – which are logically inconsistent under any contextually specified interpretation of their non-logical vocabulary. These sentences, surprisingly, turn out to largely match sentences we intuitively judge to be ungrammatical. He then argues for a four-property account of the distinction between logical and non-logical vocabulary. Functional vocabulary, and in particular logical vocabulary, is typically of a high semantic type, based in inferential patterns, works in ways identifiable across a broad class – sometimes even a universal class – of languages, and is permutation invariant. So this exciting – in Chierchia's words, game-changing – account of natural language ungrammaticality draws heavily on the semantic tradition, starting with Tarski's account of logical consequence and, in particular, the model-theoretic distinction between logical and non-logical expressions.

Our concluding paper, by Abrusán, Asher, and Van de Cruys, raises issues with Gajewski's view. Using methods of distributional semantics, they find problem cases which suggest that the logical/non-logical distinction does not work cleanly, at least for natural language. They thus reject the "logical falsity" explanation of ungrammaticality that Gajewski offers. Instead, they hold that facts about semantic composition in context explain ungrammatical unacceptability. They illustrate this by drawing on the examples of weak island ungrammaticality such as 'How tall isn't John?' and exceptives 'Some boys but John smoke'. Their essay neatly interfaces with both Glanzberg and Chierchia's discussion, adding important examples and exceptions that should be carefully considered in looking at the relationship between logic and natural language. Especially when doing so through the lens of the semantic conception.

These papers jointly show how lively the semantic tradition remains. In interpreting natural language, analyzing the difference between logic and mathematics, and getting to the bottom of what follows from what, the semantic tradition provides an important starting place for many philosophers, logicians, and mathematicians. Moreover, the connections to issues of definability, paradox, and grammaticality continue to draw interest to the fundamental issue that Tarski analyzed in 1936. We hope that these papers help to fuel a continued interest in this topic by displaying explicitly exactly how fruitful the semantic tradition can be.