



# COGNITIO

Revista de Filosofia  
Centro de Estudos de Pragmatismo

São Paulo, v. 24, n. 1, p. 1-12, jan.-dez. 2023  
e-ISSN: 2316-5278

 <https://doi.org/10.23925/2316-5278.2023v24i1:e60571>

## False triviality of truth: how Frege-Tarskian semantics misrepresents the difficulty of determining the appropriate strategic position for assertions

### *A falsa trivialidade da verdade: como a semântica Frege-Tarskiana descaracteriza o desafio de determinar a posição estratégica apropriada para asserções*

Lucas Ribeiro Vollet\*  
luvollet@gmail.com

**Abstract:** As it stands, Frege's semantic theory does not provide sufficient general keys for interpreting assertible sentences at different levels of complexity, such as (a) counterfactual-modal-intensional sentences and (b) extensional sentences. It is possible to devise that general key by adding non-classical parameters such as possible worlds, which allow sentences with a high degree of non-extensional complexity to occupy a trivial place in Tarski's hierarchy as T-scheme-eligible substitutes. The paper argues that finding these non-false (if true) conditions under which complex propositions can be trivially treated as eligible T-scheme substitutes is not trivial. It is a challenge that requires several consistency improvements to deal with various competing extensions of the predicate "true". We will conclude that Frege-Tarski's semantic conception (enriched by non-classical presuppositions), misrepresents that challenge. It distorts the problem we face in daily practice to strengthen our assertion systems, plan successful assertion strategies, and protect our assertions from semantic value reversals. The true challenge for us is pragmatic and cannot confuse (1) complicated scientific and empirical conditions of assertiveness with (2) the rewardable scoring-conditions under which one learns to use sentences in his native language.

**Keywords:** Frege. Pragmatism. Semantics. Tarski. Truth.

**Resumo:** A teoria semântica de Frege, do modo como está formulada, não fornece chaves semânticas suficientemente universais para interpretar sentenças asseríveis em diferentes níveis de complexidade, como (a) sentenças contrafactuais-modais-intensionais e (b) sentenças extensionais. É possível dar universalidade a uma chave que sirva para os dois casos adicionando parâmetros não clássicos, como mundos possíveis, que permitem que sentenças com alto grau de complexidade não extensional ocupem um lugar trivial na hierarquia de Tarski como substitutos elegíveis para o esquema T. O artigo argumenta que não é trivial encontrar essas condições não falsas (se verdadeiras) sob as quais sentenças complexas podem ser tratadas como substitutos elegíveis do esquema T. É um desafio que requer vários ajustes de consistência para lidar com várias extensões concorrentes do predicado "verdadeiro". Concluiremos que a concepção semântica de Frege-Tarski (enriquecida por pressupostos não clássicos), distorce esse desafio. Ele distorce o problema que enfrentamos na prática diária para fortalecer nossos sistemas de asserção, planejar estratégias de asserção bem-sucedidas e proteger nossas asserções de reversões de valor semântico. O desafio para nós é pragmático e não pode confundir (1) complicadas condições científicas e empíricas de assertividade com (2) as condições recompensáveis de pontuação que estabelecem os parâmetros segundo os quais alguém aprende a usar sentenças em sua língua nativa.



Artigo está licenciado sob forma de uma licença  
Creative Commons Atribuição 4.0 Internacional.

\* Doutor em Filosofia pela  
Universidade Federal de  
Santa Catarina.

**Palavras-chave:** Frege. Pragmatismo. Semântica. Verdade. Tarski.

## 1 A journey from Frege to Tarski

The path of the legacy of Frege's work *Über Sinn und Bedeutung* (1948) still the subject of curiosity. The author has presented in this small article the whole informal apparatus for thinking about the key concepts of the philosophy of modern language. By determining an intuitive scheme for thinking about the idea of linguistic composition from functions of truth, Frege's conception gives the first technical explanation of the behavior of the concept of truth, or the predicate "is true". Although this technical characterization was not yet fully developed, the germ of its maturation had already been laid. It would later be used to characterize the kind of thinking that resides in the mind or linguistic competence of someone who is able to interpret a sentence according to a generally applicable linguistic rule or key to recursively generate interpretations in that language. Thus, he anticipated the appeal that the use of this predicate should function as a computational algorithm to guide the formulation of a map between two systems. As Carnap later put it: "Semantics contains the theory of what is usually called the meaning of expressions, and hence the studies leading to the construction of a dictionary translating the object language into the metalanguage" (CARNAP, 1948, p. 10).

The landmark date for the formal maturation of that technical position was the publication of Tarski's *The Concept of Truth in Formalized Languages*.

The equivalence [between "p" and "p is true"] holds certainly if "true" is understood in the sense of the semantical concept of truth. I believe with Tarski that this is also the sense in which the word "true" is mostly used both in everyday life and in science. However, this is a psychological or historical question, which we need not here examine further. (CARNAP, 1949, p. 121).

The ability to map values onto sentences in a language in a maximally consistent way is thus the basis of our understanding of the predicate "is true" and can be revealed, as Tarski later did, by a filter (a schema with possible substitutes) that selects appropriate sentences to be truth-apt:

We have thus arrived at a statement which can indeed be accepted as the desired general definition of truth: it is formally correct and is adequate in the sense that it implies all the equivalences of the form (3) ["p" is true if and only if p] in which "p" has been replaced by any sentence of the language L. (TARSKI, 1969, p. 65).

Tarski set himself up for the application of the predicate "truth" to purely extensional languages. He promised a detailed and axiomatic account for the coincidence between truth-functional formulas and the attribution of the predicate "is true": "A definition of truth for a fragment of English should imply all the sentences of the form 'It is true that  $\phi$  if and only if  $\phi$ ' for  $\phi$  being any sentence of the fragment in question" (HORNSTEN, 2011, p. 20).

Thus, we arrive at a semantic theory with a minimum of semantic assumptions. Only the semantic contribution of the categorial terms, inserted into the linguistic syntax as a repeatable structural minimum, is considered genuinely semantic:

Tarki's solution, though simple in conception, may not be so simple in execution. [...] the goal of achieving a persistent account of logical properties makes no sense except in the context of a theory of [...] how existing members of a category contribute, and how potential members could contribute, to the truth values of the sentences in which they occur. (ETCHEMENDY, 1999, p. 50).

The minimal semantic instruction is the part of the overall meaning guidance that cannot be absent in any other assignment strategy and can be generalized as the rule of assertive prudence: the kind of discourse in which only that which cannot be retracted is considered.

## 2 The limits of deflationism: some curiosities of the concept of truth that cannot be deflated by Frege and Tarskian calculus

Frege-Tarski's characterization necessarily has a *deflationist* aspect that did not escape the late philosophers of language. Donald Davidson recognized that the foundation of twentieth-century skepticism about the substantial properties of the concept of truth, namely deflationism – the thesis that truth is not a real, substantial property – derives from a typically Fregean reading. However, this thesis would have no independent value if it were a mere philosophical statement about the mysterious character of “truth”. As a result of Frege-Tarski's characterization, it becomes apparent that the use of the “true” predicate involves processes and operations that, in fact, produce knowledge that is nothing more than a technical expression of how a language – its generative mechanisms – can be learned.<sup>1</sup> So, the concept of “truth” is non-substantive and yet an unavoidable presence in our technical mastery of a language.

In *Truth Rehabilitated*, Davidson acknowledges that the way in which the concept of truth can be used to classify propositions is intriguing. What is curious is the following: the formal relationship of correspondence between a sentence and the foundations that would make it true are not directly linked to facts in the external world, but to the possibility of this foundation lending support to all other equally true sentences. Therefore, each language has a great deal of relative freedom in selecting the sentences that pass through its truth-apt Tarskian filter. There is no common feature of ‘truth’ that would describe its property, or the “essence” of truth. The only commonality between the concept of truth in all languages is that it is a predicate whose extension includes all sentences whose truth in that language cannot be overridden by a broader rule. By examining how the Tarskian notion of satisfaction cannot describe a theory of correspondence, he supports the conclusion that the predicate “truth” does not describe a relation to the world, but a computational relation between a language and its referential possibilities:

As long as a language has the equivalent of a first order quantificational structure and no decision method, there is no way to define truth for it except by introducing a sophisticated version of reference, what Tarski called satisfaction. Tarski's satisfiers are infinite sequences which pair the variables of a language with the entities in its ontology. The interesting work of the concept of satisfaction comes in characterizing the semantic properties of open sentences, but it turns out in the end that a closed sentence is true if and only if it is satisfied by some sequence. This may suggest that we have here the makings of a correspondence theory, but it would be a Fregean theory, since every sequence satisfies every true sentence. (DAVIDSON, 2005, p. 11).

Other curiosities of the concept of “truth” remain a challenge for the scientific treatment that Frege-Tarski initiated according to our hypothesis. “Truth” classifies sentences that can be mapped as true in exactly that condition in which they can no longer be mapped as false. This property is compatible with what Quine called the “disquotational” property. Informally speaking, indeed, enclosing a phrase in quotation marks involves a choice about the protection we want to grant the sentence in a textual

---

1 As one of the three reviewers responsible for evaluating this article aptly noted, we assume a conflation of the problems of modal theory and proof theory (along with their pragmatic implications for a theory of language learning), blurring the line that separates competing ways of treating the problem of truth and that of logical consequence: “The debate about the nature of logical consequence has traditionally divided along model-theoretic vs. proof-theoretic lines” (CARET; HJORTLAND, 2015, p. 4). In twentieth-century philosophy of language, it is a well-known fact that semantics can be studied from a model-theoretic perspective (think of Tarski's semantics for first-order logic) or with a focus on the nature of derivations and proofs (for example, Gentzen's sequent calculus). However, we do not believe that confusion is the basis of our guiding threat. These formats of study are, we take it, different angles that illuminate each other. In Davidson, for example, we see the application of Tarski's T-convention (model-theoretically) to gain insights into “how we can go from truth to something like meaning” (DAVIDSON, 2001, p. 74). Thus, representing the truth conditions of a sentence through a model provides a non-isolated vantage point capable of illuminating the problem of building an empirically testable interpretive theory for communication and language learning. In the hope of illuminating the issue from additional angles, we even include an examination of Hempel's confirmation paradoxes. The question of whether this particular syncretic fusion of themes is the most appropriate is not settled, but we can consider it a methodological choice of the present article. Nevertheless, I would like to thank the reviewer for the attention he has given to this aspect of the article.

composition. This justifies the assumption that removing it from the quotation has the opposite effect: it rescues the maximum sensitivity of the sentence to the semantic influences it can receive, that is, those influences that can contribute to reward it as true in the exact condition in which it cannot be false.

This supports the deflationist conclusion that whatever “truth” classifies is nothing but a predicate that correlates the extension of other predicates, and thus the property classified by this concept cannot be used as a gender that distinguishes things as green or blue in a language. It can only be used when one language is in a position to talk about another safely (as metalanguage). But, as Davidson himself pointed out, this feature of the “true” predicate cannot teach anyone what languages capable of decomposing their sentences into non-paradoxical positions have in common: “Disquotation gives the extension of a truth predicate for a single language; if we ask what all such predicates have in common, disquotation cannot answer” (DAVIDSON, 2005, p. 11).

In Davidson’s words: “if we want to know under what conditions a sentence containing a truth predicate is true, we cannot use that predicate in the disquotational mode” (DAVIDSON, 2015, p. 11). Davidson is concerned that the filter for sentences that can be disquoted does not provide relevant information about the profile of the language that has the expressive power to present it as disquoted. While this criterion shows the extension of the predicate “it is true” for a language, it does not show what distinguishes that truth-extension from, say, a truth-extension of a language that cannot be learned with the same ease (for example, the technical language required to learn theoretical physics). Moreover, it is not possible to use the idea of “disquoting” alone to determine the difference between a successful assertive strategy, that derives its expressive power from semantically coherent language; and an unsuccessful strategy, that does not derive its power from Tarskian semantic conditions and therefore can be enforced under conditions under which it could be overridden – defeated. It cannot show how the predicate “is true” represents the *specification* of its assertibility as true only under conditions in which it is not false.

The lesson of this chapter is that even Tarski could not maintain his conception of truth without strengthening it to excluding surrogate-sentences for the T-scheme that – although *syntactically impeccable* – do not meet more stringent specification conditions. Through Davidson’s word, we also discover that there is something fundamentally non-trivial about a culture’s use of the word “it is true”: “Tarski’s truth definitions are not trivial, and they reveal something deep about languages of any serious expressive power” (DAVIDSON, 2001, p.11). This is easily explained. What *really* excludes “falsehood” from the extension of the predicate “true” must satisfy a stronger requirement in order to explain what the “true” predicate specifies. This stronger requirement must be described by an ideational or normative character, otherwise the guarantees that one rule does not override another would be too weak. Weak conditions do not give good strategic knowledge of the limits and reach of our assertions. Our assertion strategies, for example, to say that “Russia is not a communist country”, would not be linked to a theoretical understanding of the conditions under which this is “not false” under our premises. Thus, the coherence of our use of the true predicate is linked to the condition of language norms to provide a stable environment for choosing non-defeatist assertion strategies.

### 3 Some difficulties disguised as trivial: the limits of the Frege-Tarskian conception

Tarski’s lesson is that the semantic content unlocked by the rule or model for distinguishing the truth of a sentence cannot adequately be described without inconsistency, unless the language is broken into secure platforms, each one protected against the content described in the lower one. “Tarski has taught us that this description will, on pain of contradiction, have to be given in a more encompassing framework than the language for which the models are intended” (HORSTEN, 2011, p. 21).

This lesson contributes to thinking of the semantic content that accompanies the description of “truth” as unnecessary and even harmful and dangerous. For wherever we need a rule or model to describe truth, either a meta-language is required, or we are vulnerable to contradiction. However, this turns the first question into a new one, as the question arises about the truth of metalanguage sentences. Thus, a new semantic model has to be worked out in a higher-level language. Circular regress is inevitable and, therefore, many philosophers take the Tarskian lesson as a reason for a semantic deflationism, by theorizing that truth is not a genuine property. But we must still make sense of our inevitable truth talk. Instead of looking for the semantic identity of sentences with the same truth condition, which would lead to multiplying languages, we simply determine these sentences according to the same rules: “The axiomatic approach does not suffer from the regress problem: An axiomatic truth theory at least partially gives the meaning of the truth predicate of the language in which it is stated” (HORSTEN, 2011, p. 21).

In summary, sentences whose assertion can be made under the same conditions as saying that *it is true* do not produce inconsistencies, are *avored* to be asserted. That is, those sentences which are rewarded with the predicate “true” under conditions where this predicate cannot be reversed into falsehood, are those sentences that find themselves in a privileged condition: they are the theorems of a theory of truth, or the sentences that cannot be revised in a theoretical construction of truth for a system. Tarski thought that this would place these sentences in arbitrary positions. In particular, he suspected that any (consistent) collection of axioms of truth would have an “arbitrary and accidental” character (HORSTEN, 2011, p. 22).

In this framework, certain sentences that would be true in one scientific system could be false in another. This may have caused Tarski some distress, but this consequence was not taken as undesired by the tradition’s reception of his theory:

Tarski wanted to define truth. He saw that truth could not be defined in the object language: It has to be done in an essentially stronger metalanguage. In his definition, Tarski assigned a pivotal role to the Tarski-biconditionals. Nowadays, many a deflationist philosopher takes a suitable class of Tarski-biconditionals as the axioms of her theory of truth. (HORSTEN, 2011, p. 47).

In a perfect ideal condition, Tarskian requirements prevail: there will not be any relevant circumstances for a semantic theory that can distinguish a true sentence from any other true sentence. The condition in which they are non-false, thus, is one and the same – at least for semantic purposes. Everything looks trivial as far as “truth” is concerned, and we can slingshot it around. Indeed, semantic purposes seem to be just the ones that matter when we idealize the conditions of interaction and communication. The predicate “truth” has the function of classifying sentences that are sensitive to the same type of norm, that is, that are corrected by the same criteria, and the knowledge of these criteria is not a metaphysical knowledge of “true things (propositions)”. It is the mere knowledge of the primitive regularities of a linguistic structure, insofar as they represent the projection of semantic values, those values that contribute to the mapping relation of the sentence and what the sentence represents in cumulative (non-reversible) conditions. But we have not yet answered the most important question: What is this knowledge that we have when we are in this state? According to Tarski, it is the knowledge we have when we know how to replace the T-scheme with legitimate candidates. We will challenge this thesis below.

The criticism to which we shall refer in this chapter is the following. Frege-Tarski’s characterization highlights an aspect of our interpretive practices that would supposedly give us a decoding key for interpreting propositions by the minimum that can be learned from their assertion, i.e., the element that is projected by the proposition’s assertion and cannot be overridden or withdrawn. This is its similarity to all true propositions (if true) and its similarity to all false propositions (if false). The problem is that there are no known or controlled limits that discipline our ability to generate sentences under these conditions. The similarity of a true sentence to all other true sentences can be generated by the consistency of



a predicate that defines the extension of “incompatibility of p and non-p”. But this predicate can be generated and programmed into many different scientific systems, each of which has its own extensional version of this incompatibility. So, we are dealing with the possibility of limitless discretion to generate the extension of “is true”.

We can formulate this problem in another way. We are not empiricists, but an empiricist’s perspective can help understand what is scandalous about the Frege-Tarskian conception. An inveterate empiricist might protest that, according to this conception, any language can have an “is true” predicate as long as it has sufficient encoding facilities to give consistency to the antagonism between extension and anti-extension of the truth predicate. Several solutions to this problem can be given, as fixed-point theories. Complex intensional and counterfactual content within sentences can then be tamed in that language. The scandalous point is that even the liar paradox can be tamed in this way, as Kripke attempted to prove:

Intuitively, the situation seems to be as follows. Although the smallest fixed point is probably the most natural model for the intuitive concept of truth, and is the model generated by our instructions to the imaginary subject, the other fixed points never conflict with these instructions. We could consistently use the word; “true” so as to give a truth value to such a sentence as (3 [(3) is true]) without violating the idea that a sentence should be asserted to be true precisely when we would assert the sentence itself. The same does not hold for the paradoxical sentences. (KRIPKE, 1975, p. 708).

Enriching the language to the point where it can disquote even the liar paradox is feasible. Some adjustments will be demanded – not everything is for free – but still, the liar will talk and walk as everyone else. The price of taming this kind of complexity is that we lose some of our normative reference to nonsense. No sentence will be so airtight that it cannot be given an innocuous place in language.

Certain states of peace are not always desirable. Learning a language will hardly be easier for a child if it combines absurdities under the same umbrella of meaning. When the difference between absurdity and meaning becomes so much less important and normative that even sentences with contradictory information are possible, we must learn a language by expanding our notion of consistency-which is hardly the most appropriate case for a child.

For the empiricist the liar is not even the most dangerous problem. This limitless enrichment is arbitrary because it gives homogeneous treatment to a complex sentence like (+t) “World War II repeated elements of the Hundred Years War” and a much simpler one like (t) “World War II ended in 1945”. Children can learn both on the same syntactic level when they are put on the same level of *disquoting literalism*. But the second has a much more innocuous relationship to truth than the first, and this – for the empiricist – must be reflected 1. in the difficulties involved in showing data that counts as evidence for it to be proved and 2. in the difficulties involved in selecting the data that teaches the *habits* of use of complex sentences, forming a complex general belief like empirical laws. It is not to be expected that a child knows complex statements about the second war just as it learns to generate syntactical compositions of his/her language.

The scandal is that by the Frege-Tarskian criterion, both +t and t sentences can be equally disquoted as possible substitutes of p in the famous scheme “‘p’ is true if and only if p”. As Michael Dummett has well recognized, we can always give a straightforward characterization of the relative conditions of assertability: “In a semantics of this latter kind, relativized truth represents either truth in a world, (...) or else assertability in a state of information” (DUMMETT, 1991, p. 45). We can indeed, but not without costs. This paper was written under the assumption that, although the T-scheme represents the ability to present any relative statement in absolute conditions as they are converted to T-schemes, it is unable to chart the costs that different T-scheme’s conversions have for our learning ability and our habits of language use. Since the data one should use to reward a child’s learning is not even available for general complex sentences, it is hardly easy to identify what it means to say that “+t is true if +t”.

## 4 The false triviality of truth-conditions

Tarski's theory makes it seem that straightforward interpretations of complex modal and counterfactual statements is cheap or even gratuitous. This would be good business for liars and cheats: confusing and hermetic sentences would get as much credit as simple ones. The Frege-Tarskian argument inevitably leads us to regard the predicate "truth" as a kind of instrument of an undisciplined and unbounded project of "proof-simplification". Even the most dependent, contingent, and difficult to prove statement can be characterized by a simple rule connecting it with a possible value. Under these premises, we use this predicate when we want to express our discretionary power, and give absolute-straightforward interpretations, even for complex ungrounded sentences that require a lot of empirical and experimental climbing to prove. Triviality would be tangible for any sentence, no matter how complex, hermetic, hypothetical, etc.

We think that liars and cheaters still have no reason to celebrate. We argue that the ambition to find those unified criteria to judge meaning is all an illusion, probably inspired by the assumption that every sentence can be modeled in a way that makes it look like a sentence with a direct correlation to its minimum level of dependency – that is, if it has a truth value in the smallest fixed point. This would lead to the highly desirable situation in which every assignment of truth to a proposition corresponds to its criterion of proof, replacing the empiricist myth of "given" and "confirmation" with the computational myth of the "minimal point of dependence". But what has changed with these artificial devices for determining truth value (at the irreversible point) is not our ability to avoid paradoxes like Hempel's in 1945<sup>2</sup>. (where the evidence for any non-black non-crow would provide potential evidence for the claim that all crows are black). The extensional equivalence between "All crows are black" and "All non-crows are not black" further leads to the troubling situation that proving a general statement is much more difficult than proving a particular statement. So, what do we gain from Tarski and Kripke teaching us to give an innocuous place to ungrounded propositions? We certainly have not lightened the lives of scientists to prove that all crows are black. This strategy has merely improved our ability to recognize confusion and paradox in an intelligible way by providing us with a basis for predicting what the liar is trying to say, or how many linguistic rules he needs to neutralize in his mind so that "I lie" is encoded in a familiar way. The familiar way would be to project the statement's truth only to the point where its falsity is no longer in question. But if one gives a reason to a madman to make his statements more publicly acceptable and understandable, that does not render the madman reasonable. This is, at best, an improvised means of making his madness predictable.

Let's look at the problem from a pragmatic perspective. In any serious scientific solution of the problem of determining the matching correspondences between equivalent propositions, we need a codification of the *favorability* (or something like the "truth-priority") of the assertion under possible conditions. Dummett noted that: "the use of a quasi-ordering between the statement values is far superior, for the purposes of logic, to that of a classification of them into those that are and are not designated" (DUMMETT, 1991, p. 44). However, if we take the trouble to order our system of truths, we proceed on terrain that is hardly conducive to a trivial semantic classification of truths. Because this involves a notion of "proof" whose extension is not interchangeable with the notion of "truth". As truth is ordered, there will be different ways of satisfying the predicate (is true) and related different notions of consistency. Truth and proof can be artificially equated in any mathematical construction, but in real debates one cannot simply avoid meeting stronger requirements for proof (those of physics). Furthermore, standardizing it to conform to a universal scheme as simple as the one required for children to generate sentences from their language (the T-scheme) cannot be trivial. This means that even if all

---

2 Hempel prevented progress in developing a logic of confirmation by demonstrating his triviality result (HEMPEL, 1945) that any statement (observational report) E is consistent with any statement (hypothesis) given some basic conditions for the adequacy of any confirmatory relation.

native speakers of a language endorse sentences  $p$  and  $q$  under conditions  $t$ , the sentences may depend on more or fewer preconditions in order to be proved, and thus a single rule correlating truth and proof cannot determine that sentence's use (making the job of linguist-translators much more difficult).

Tarski once complained that in mathematics the notions of truth and proof are not in conflict: "there is no conflict between the notion of proof and truth in mathematics" (TARSKI, 1969, p. 77). In a mute tone he seems to regret that this is not the case in natural language and empirical science, and we may even suspect that his whole theory – and perhaps the mission of his intellectual life – is to attempt to bestow upon other realms of expression the same expressive and normative blessings inherent in mathematical constructs. The author should have guessed that mathematics is the terrain where arbitrariness knows no bounds. In our practice of argumentation and rational disputes, we cannot assume that we can prove with the same ease the causality of a drug and the cure of a disease, as well as any statement about the accurate date and spatial coordinate of an event.

## 5 Winning conditions for assertability and strong specifications of Meaning: the non-overridable nature of the truth under strong expressive conditions

What is trivial in minimalist and deflationary explanations of true is the way in which sentences classified as true by this system appear to speakers of the language: as sentences that can be substituted without difficulty into Tarski's T-scheme, claiming nothing more than their classification as true – their assignment in the anti-extension of "is false" – is required for them to be true. Of course, triviality here is an illusion of perspective. It is not so easy to find an assertion strategy in which only the minimum is taken as a proof-parameter, i.e., only the condition under which the proposition cannot be false if it is true. For example, it is not easy to decide under what conditions the sentence "The moon is a round satellite" is rewarded as correct under the same conditions under which "The moon circles the Earth". There are a lot of assertion conditions rewarding one but not the other, others rewarding both, others still, rewarding none. Thus, it is not easy to say what the truth or falsity of these sentences has in common, since they cannot be learned according to the same scoring-rewarding system. Since the conditions under which a sentence is proved are not always simple, learning all sentences by learning the extension of "is true" can be anything but trivial.

Several subterranean steps must be taken for truth and proof to coincide. Sometimes the steps are artificial: they serve only the dubious purpose of making liars less threatening. And it gets worse for modal and counterfactual propositions. The predicate "true" is supposed to eliminate any proof-underspecification of the sentence that would make that sentence vulnerable to being false in the same conditions in which it was asserted as true. But as we've seen, different languages and scientific systems may identify different extensions for the incompatibility of  $p$  and non- $p$ . To develop the frame for deepening the ongoing proof-specification of any assertive strategy, we need to engage with some conceptual knowledge that would ask for stronger equivalence conditions. Mere extensional conditions will require a stronger metalanguage. So, we cannot say that we know the truth if we have only a weak extensional knowledge of the equivalences of languages. We need a rule to eliminate the degree of proof-underspecification that each sign has. And we need to idealize the conceptual conditions for specific instructions of proof and meaning.

To test this argument, we can describe an imaginary situation. In a situation where there is a very idealized consensus and a strong scoring-rewarding system to decide what is a valid and what is an invalid assertive move, there is also a consensus on how much a sentence can take advantage of each instance that confirms or refutes it, turning that confirmation (refutation) into a single unit of semantic value transmitting cumulative credibility. Those are the circumstances where truth and proof coincide



trivially. All seems trivial in the correlation structures, and then it is trivial to use any mathematical technique to describe ideal or hypothetical correlations between those structures. This is indeed a very rare situation if it is possible at all. In this particular imaginary situation, the only way to violate this condition is to open margins of semantic divergence large enough to turn the assertion into a fraud at worst or a self-defeating move at best. To speak of a semantic condition is nothing but to speak of the *game condition* in which a move is not elastic or complicated to the point of being expressed as a trick or an exception rule, that would be rewarded differently, like the scoring card.

What is not trivial here is exactly what the predicate “true” is supposed to represent: the strengthening of the correctness of the interpretation in a game. Because this game is not a mere joke. It does not describe a triviality. It describes a complex state of stable human practices that depend on a stable notion of “consistency” in order to charge reasons and compromises. The stability of our notion of consistency is also not that easy to achieve. We need to protect our languages with strong normative requirements to achieve a notion of consistency for it that is workable or usable in human practice, as Dummett has noted: “we must impose certain conditions on the total set of logical laws we are stipulating to govern a logical constant. The first of these was harmony, a stronger requirement than consistency, but one which guarantees it” (DUMMETT, 1991, p. 246).

The expressive power of language is normative because the capacities to perceive and engage structural aspects of language as possibilities of assertion have correctness conditions to guide possible specifications of meaning without paradox. The supposed freedom to create interpretations for propositions by systematic assignments has some normative constraints, but not arbitrary constraints. Not every kind of constraint is appropriate. Sometimes a bad scoring system can bring too many “scoring-cards” into play. Of course, sometimes we have to “teach” language to deal with antinomies or sentences that cancel each other out by adding codes. Some scoring-cards may be included to help communication. The kind of syntactic-grammatical constraints that any language imposes to specify its meaning statements in a coherent way or without categorical violations derives from the knowledge of the possible solutions to misunderstandings, antinomies and theoretical disputes that this language can deal with. As we evolve in conflict to new problems, our language should be able to evolve its categorial and syntactic possibilities as well in the face of new antinomies. In a more precise and technical statement, the word “true” is precisely what characterizes the theoretical understanding that speakers of a language have of these constraints and their ability to arbitrate solutions to these disputes and disagreements, but it must also include our ability to find room for solutions to new antinomies or type-conflicts.

## 6 Final considerations

The classical Fregean ambitions presupposed an untenable unified-ideal of linguistic competence for sentences of different levels of complexity. It would fail to provide a universal key valid for unlocking semantic values of both intensional and extensional sentences. These ideal assumptions were unable to respond to substitutional problems with intensional, modal, and counterfactual content, eventually leading to an explosion of non-classical logics. These, in turn, retained Frege’s assumption and merely extended it to sentences with more complex extensions (possible worlds) or no-truth-values (many-value-logic). The new assumption would be: no matter how modally and counterfactually complex a sentence is, it could simply take a place as a substitute for  $p$  in Tarski’s scheme. Even paradox and ungrounded sentences would have a place in this hierarchy. And so, with some degree of magic, a modal sentence could be interpreted as true under as much straightforward proof conditions as sentences without modal conditions. All it takes is for the predicate “is true” to be antagonistic to its anti-extension, and for that it is enough to enrich the language with platforms and fixed points in such a way that complex modal sentences can have a single interpretation in it – they cannot be turned into false if they are true.

As we have seen in this article, this assumption needs to be revised to answer to some problems. One problem is that the unrestricted correlation between the notion of truth and the notion of proof is possible only if we strengthen the language so that it accepts only non-defeasible propositions as T-scheme substitutes. Hypothesis, speculation, modal and intensional content must therefore at least meet this criterion. Only non-defeatist strategies for asserting these propositions are sufficient. We may of course follow non-classic assumptions and find possible-worldly parameters to know a rule for asserting intensional-modal sentences. But not without strengthening language so that all sentences, no matter how complex their world selection, are described as absolutely true or false by the same unambiguous conditions. By strengthening it, however, we pay the price in terms of certain normative and conceptual adjustments of categories or logical constants to find the conditions of favorability under which intensional and extensional sentences can be asserted in a non-defeatist way. Only under these conditions will the programming of a proof of the sentence (whether of “a” or “b” stage of complexity) be consistent with the conditions under which that sentence is not false if it is true. The lesson is that this raises a new problem: knowledge of the point of favorability at which a sentence’s assertion is irreversible is no longer just knowledge about the substitutability of T-schemes (and the syntactic features of legitimate candidates). It is a broader normative knowledge that we acquire as we improve our codes for resolving antinomies and protecting our sentences from defeatist interpretations.

It is fair to say, then, that the way Frege-Tarski formulated the problem does not present the real challenge involved in determining the meaning of a sentence. There is more “logic” (in the sense of more rules) in this process than their premises seem to allow. It is definitely possible to find the point at which a sentence is correctly used in a language and simultaneously classified as true, false (or having the value “none”). However, this comes with another pragmatic challenge. We cannot learn to determine this key point of sentence sensitivity by simply inserting the sentence into the T-scheme. Since the conditions under which we can say that an interpretation is correct are more stringent than those under which we can say that it is *assignable* to a reference, this adds a requirement to the understanding of the predicate “true” that can only be satisfied by a comprehensive knowledge of its ability to leverage non-defeatist assertive strategies. Extending these conditions to understanding liars is of limited value: we can predict the liar’s claims without having to violate our concept of rationality. But we have also forgotten – or can no longer understand – the difference between what is unsuccessful and self-defeating in the sentences of the liar and the deceiver.

Now we have described a highly non-trivial condition. The requirements for the use of the truth predicate have become so stringent that we can no longer imagine that any syntactically well-formed sentence can be considered a substitute for the T-scheme without further difficulty. The non-triviality of semantics is related to the challenge of distinguishing languages whose structure does not undermine our ability to make non-trivial claims, as hypothetical and intensional sentences. This knowledge cannot be called trivial, because it must select consistency rules that allow us to navigate the universe of possible-world-states without reversing our sentences from true to false and from false to true. It is the challenge of selecting consistency-rules working even to assert sentences with modal and counterfactual contents, that do not have a single condition for falsehood. The language of law, for example, is always faced with the challenge of improving its consistency codes.

Ultimately, our language appears not as a natural system of rules, but as a protective tool that upkeeps successful claims and provides the strategic basis for the cumulative (irreversible) advancement of our truth claims. But the ideal state of stability we can achieve to produce only non-defeatist assertions cannot, for a more serious reason, be mere knowledge of the properties required for eligible well-formed substitutes of the T-scheme. Each time we recode our language, we increase the extension of “incompatibility between p and non-p”, so that our notion of consistency tends to expand as we learn sentences from theoretical physics, sociology, and so on. David Hume would say that we extend our habit for learning general propositions; Immanuel Kant would say that the general content of these

sentences is linked to the conditions of the possibility of human judgment. We do not need to go back that far in the discussion. But we can know right now that we cannot expect these learning conditions to be as simple as those required for generating keys to produce sentences in natural languages.

## References


- CARET, C. R. & HJORTLAND, O. T. Logical Consequence: Its nature, structure, and application. *In: CARET, Colin R.; HJORTLAND, Ole T. (eds.). Foundations of Logical Consequence*. Reino Unido: Oxford University Press, 2015.
- CARNAP, R. *Introduction to semantics*. Cambridge, MA: Harvard University Press, 1948.
- CARNAP, R. Truth and confirmation. *In: FIEGL, H.; SELLARS, W. (Eds.). Readings in philosophical analysis*. New York: Appleton-Century-Crofts, 1949. p. 119-127.
- DAVIDSON, D. *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press, 2001.
- DAVIDSON, D. *Truth, Language and History*. Oxford: Clarendon Press; New York: Oxford University Press, 2005.
- DUMMETT, M. *The Logical Basis of Metaphysics*. Cambridge, MA: Harvard University Press, 1991.
- ETCHEMENDY, J. *The Concept of Logical Consequence*. Stanford, CA: Center for the Study of Language and Inf. 1999.
- FREGE, G. Sense and Reference. *The Philosophical Review*, v. 57, n. 3, p. 209–230, 1948. [<https://doi.org/10.2307/2181485>].
- HEMPEL, C. G. Studies in the Logic of Confirmation I. *Mind*, 54(13), p. 1-26, 1945. [<https://doi.org/10.1093/mind/LIV.213.1>].
- HORSTEN, L. *The Tarskian Turn: Deflationism and Axiomatic Truth*. Cambridge, MA: MIT Press, 2011.
- KRIPKE, S. Outline of a Theory of Truth. *The Journal of Philosophy. Seventy-Second Annual Meeting American Philosophical Association, Eastern Division*, v. 72, n. 19, p. 690-716, 1975. [<https://doi.org/10.2307/2024634>].
- TARSKI, A. Truth and proof. *Scientific American*, 220(6), p. 63-77, 1969. [<https://doi.org/10.1038/scientificamerican0669-63>].



# COGNITIO

Revista de Filosofia  
Centro de Estudos de Pragmatismo

São Paulo, v. 24, n. 1, p. 1-12, jan.-dez. 2023  
e-ISSN: 2316-5278

 <https://doi.org/10.23925/2316-5278.2023v24i1:e60571>