SOME PROBLEMS REGARDING
THE DISQUOTATIONAL PERSPECTIVE ON TRUTH

Abstract

According to the proponents of disquotationalism, „being true” is „strongly” equivalent to „being a device of (sentential) disquotation”. This means, on the one hand, that the disquotational use of the truth predicate covers everything there is to say about truth, and on the other hand, that the disquotational axioms are warranted. I am expanding these issues by pointing out the main features of the disquotational theories of truth, their formation from Tarski’s „concept of truth” (on the basis of an elementary intuition about truth present from Aristotle’s philosophy on), I give an interpretation of the way they build their truth predicate and I present some arguments and counter-arguments for or against the deflated conception of truth, based on the requirement of analyticity of the T-sentences.

Key words: formal semantics, deflationism, disquotational truth theory, T-sentence, Tarski

Disquotational Truth Theories. General Features

When conceiving or dealing with a philosophical theory there are a few things that should be investigated. First, one has to know what the theory is about. This means, of course, knowing what its basic concepts and premises are, and also what they are supposed to lead to. Secondly, it is important to understand how the concepts direct the theory in the respective direction – which includes proving whether the basic elements of the theory were precise and strong enough to let the theory do its purported job. If the system is alright (to the extent to which a philosophical system can be alright), it should be important to see what effect it had over the general realm of philosophy – if and how the other main philosophical concepts fit into its schema, or what is left of them, if they don’t.

The disquotational theories are a branch of the deflationary theories of truth – and as such, their endeavours revolve around the concept of truth (and the

1 In the following, I will use italic characters for emphasis (and, in disquotational structures, for expressions translated in the metalanguage), simple quotation marks (‘’) to brand a word used in a loose sense, double quotation marks („”) for citations or other mentioned (as opposed to used) expressions, and straight quotation marks («») for quotes within quotes.
The disquotational perspective of truth

The truth predicate is their basic object of research and the T-sentences are their main axioms (whether conceived as the – quite large – set of instances of „T(«p») if p”, or as consequences of a principle according to which the T-sentences are – somehow – analytic).

The purpose of disquotationalism is, mainly, to purge the truth predicate of its cumbersome metaphysical weight, of suspicious obscurities and paradoxes, and to reconstruct it with the help of a formal theory, all the while keeping the most important intuitions connected to it intact. This relatively new path in logical semantics has been taken because of the shortcomings of its main competitors, the correspondence theories of truth.

In spite of their tempting obviousness, many of these theories suffer from the same issue which pains most of the obvious theories: they are conceptually obscure and leave plenty of place for inexactitudes, including the worst type of them, the paradoxes. One of the strategies used to remedy these inexactitudes is the deflation of the concept of truth; this strategy is worthy of a further investigation, first of all, because of the problems issued by the truth, if considered as a deeply metaphysical concept, and secondly due to a simple discovery: it might actually happen that „is true” be employed solely as a tool for disquotation, useful, for instance, in order to simplify sentences uttered on the metacomunicative level of daily conversations. Since it was already present in Aristotele’s definition of truth², one should be able to say that the disquotational feature is at least in part integrated in the intuitive background of the classical concept of truth. And then, of course, this is also the reason why the proponents of disquotationalism use this predicate for their semantic theories, otherwise they could have invented a new concept.

Concerning the internal mechanisms of the disquotational truth theories, the story of truth-as-a-device-for-disquotation could be interpreted³ as unfolding like this. We have a name for a term or for a sentence, and now we want to know what the name was about. If the naming process has been worked out properly⁴, the step seems to be straightforward – if there are no two objects that have received the same name, there should be some sort of function (like „naming”) which sends names back to their original reference. In the case of terms, the function is, of

---

² To say of what is that it is, or of what is not that it is not, is true.
³ The disquotational theory doesn’t actually introduce the truth predicate in order to fix the reference of sentence names. The main problem is how to use and understand the truth predicate of a language, and the conclusion is that truth can only be understood as a disquotational device. What follows is therefore an interpretation of the disquotational story.
⁴ No object should receive two different names and no two objects should receive the same name. Also, it would be quite nice if there were no names that don’t refer to anything.
course, the referential function: „The reference of «D» is D”\(^5\). For sentences, however, the reference is not suitable for the disquotational job. At this point, the disquotationalist notices that we have a tool for this case as well, and that would be the truth predicate, an operator that behaves like the inverse of the naming function – thus, to determine what it is for a name of a sentence to fall under the incidence of the predicate „T” is the same as to determine the value taken by „naming\(^{-1}\)” when applied to it. For sentences, the truth predicate is the only possible device of disquotation\(^6\), and the only function of the truth predicate is to „disquote” sentence-names\(^7\): in order to determine what a name of a sentence (or set of sentences) stands for, it is enough to apply to it the truth predicate and one will get immediately back to the „original sentence”. What allows one to get back that ‘fast’ to the original sentence is exactly the fact that the disquotational feature is all there is to say about the truth predicate. It is all that is left of its intuitive background.

For the disquotationalist, truth cannot be considered a genuine concept. This fact is self-evident, seeing that the disquotationalist truth, or the predicate „true” is defined as nothing more than a device of disquotation. Technically speaking, „...is true” does behave like a formal predicate, and the deflationists do regard it as such. But since truth cannot keep its meaning as a concept anymore, in the same way it will lose its attributive power, hence its authenticity as a genuine predicate. When we say about snow that it be white, we emphasise a property of its nature, the whiteness. From a rhetorical or communicative point of view\(^8\), truth is also used as a device for emphasis. However, from an epistemological perspective, „is true” doesn't put forward anything that the original sentence wouldn't put forward itself. This is why the discussed notion differs from any other predicate – except, perhaps, from „exists”, or better, „refers to”. Saying that „«Snow» refers to snow” emphasises as much about snow as saying that „«Snow is white» is true” emphasises about „Snow is white”. It is possible that the correspondential theories of truth be silently based on the idea that the addition of „is true” functions like a finger pointed to the real states of affairs in question. In fact, „is true” is still nothing but a part of the language, and as such it doesn’t bring the sentences it is applied to any closer to „reality”\(^5\).

\(^5\) Or, properly stated, „<D> refers to and only to D, if <D> refers at all or if D exists” (Volker Halbach, *Semantics and Deflationism*, unpublished Habilitationsschrift, 2001, p 131).

\(^6\) One could think of other disquotationalist-like devices (*is assertable* or *is worthy to be believed* for instance), but they all seem to derive somehow from the concept of truth.

\(^7\) „It is a fundamental intuition about truth that from any sentence A the inference to another sentence that asserts that A is true is warranted. And conversely: from the latter sentence the inference to A is warranted.” (Anil Gupta, „Truth and Paradox”, *Journal of Philosophical Logic*, vol. 11, 1982, p. 6).

\(^8\) Almost every example used to explicitate the prosentential theories of truth introduces the discussed predicate by means of a dialogue.
There are two main problems that have to be considered when pondering whether our intuitions about truth can be completely reduced to the disquotationalist view. For an affirmative answer, one has to check first if all our natural intuitions about truth are covered by the T-sentences. This issue has been already touched upon in the previous paragraph. The second problem is of equal importance: are all our formal intuitions about truth covered by the T-sentences? We could rephrase the last question like this: does the disquotational concept of truth solve the paradoxes about truth? Do the deflationists manage to construct a system that reaches its purposes?

A T-sentence expresses the simple fact that its right side is equivalent to its left side, in other words, that calling a sentence „true” doesn’t bring forth something essentially new compared to its simple enunciation. What one would like to know now is the sense in which the right and the left side of the disquotation sentences are equivalent. Basically this problem directs the discussion toward the „degree of necessity” under which these sentences are considered to be equivalent: is the sentence „Snow is white” formally, or materially equivalent to affirmation of its truth?

Tarski

The talk of a „left” and a „right” hand side of a disquotation-type sentence was initiated by Tarski’s definition of truth, whose adequacy criterion was given explicitly in the form of what was to become, a few decades later, the structure of the disquotation sentences themselves. Tarski didn’t call his material adequacy criterion disquotational, nor is it clear that he thought of it in the same way as the proponents of the disquotation sentences do. But his famous „snow is white” equivalence, nowadays the most typical example of a T-sentence, was the first one to have adopted this shape explicitly.

The intuition that led to this idea came up before Tarski. According to Gupta and Belnap, it seems that a trace of what was to become the disquotation theory of truth could be found even in Aristotle’s Categories: „…there being a man reciprocates as to implication of existence with the true statement about it: if there is a man, the statement whereby we say that there is a man is true, and re-

---

9 In the words of Paul Horwich (Truth, Basil Blackwell, Oxford, 1990), „most of the interesting facts to be explained concern relations between truth and certain other matters; and in such cases it is perfectly proper to make use of theories about these other matters, and not to explain that all the explanatory work be done by the theory of truth in isolation. […] contrary to what has been generally presupposed, the notion of truth is not even involved in the problem” (p. 7-8).

crăiţa FLORESCU

ciprocally – since if the statement whereby we say that there is a man is true, there is a man”11. In the interpretation of Gupta and Belnap, this already brings forth the first observation that leads to the constitution of the disquotation sentences: the analyticity of the equivalence between there being a man and the truth of „There is a man”. As a second observation, this equivalence is also exhaustive with regard to the concept of truth. In today’s terms, truth is (necessity) nothing but (sufficiency) a device of disquotation. From Gupta and Belnap’s point of view, this observation came along together with Ramsey’s theory of truth.

Tarski hasn’t committed himself explicitly to deflationism, as a matter of fact he was quite fond of the correspondence theory of truth12. However, he did notice the problems created by the ‘naïve’ definitions of truth and, trying to root them out, he moved smoothly away from what he explicitly adhered to13. It is important to stress the fact that, in Tarski’s theory of truth, the part that really „does justice to the intuitions expressed by Aristotle’s words” was not the definition of truth, but only its material adequacy principle. Tarski’s recursive definition of truth was based on the concept of satisfaction, defined exclusively with the help of Tarski’s axiomatic apparatus which was responsible for the formal correctness of the system. Further on, it is interesting to see the way Tarski came back to his material condition: „the question now arises whether this definition […] is also materially correct – at least in the sense previously laid down in the convention T. It can be shown that the answer to this question is affirmative: Def. 23 [the definition of truth] is an adequate definition of truth in the sense of convention T, since its consequences include all those required by this convention”. Until now nothing surprising. But Tarski goes on: „Nevertheless it can be seen without difficulty (from the fact that the number of these consequences is infinite) that the exact and general establishment of this fact has no place within the limits of the considera-

11 Quoted from Gupta and Belnap, op. cit., p. 1., my emphasis.
12 „We should like our definition [of the term «true»] to do justice to the intuitions which adhere to the classical Aristotelian conception of truth – intuitions which find their expression in the well-known words of Aristotle’ Metaphysics […]. If we wished to adapt ourselves to modern philosophical terminology, we could perhaps express this conception by means of the familiar formula: The truth of a sentence consists in its agreement with (or correspondence to) reality. […] If we agree to speak of the designata of sentences as „states of affairs”, we could possibly use for the same purpose the following phrase: A sentence is true if it designates an existing state of affairs.” (Alfred Tarski, „The Semantic Conception of Truth and the Foundations of Semantics”, in Steven R. Givant, Ralph N. McKenzie (eds.), Alfred Tarski, Collected Papers, Volume 2 (1935 – 1944), Birkhäuser, Basel, Boston, Stuttgart, pp. 666–667).
13 Here is the immediate sequel of the above-quoted paragraph: „However, all these formulations can lead to various misunderstandings, for none of them is sufficiently precise and clear […]; at any rate, none of them can be considered a satisfactory definition of truth. It is up to us to look for a more precise expression of our intuitions” (Alfred Tarski, op. cit., p. 667).
The disquotational perspective of truth

tions so far brought forward. The proof would require the setting up of an entirely new apparatus: in fact it involves the transition to a level one step higher – to the meta-metalanguage, which would have to be preceded by the formalization of the metatheory which forms the foundation of our investigations\textsuperscript{14}. Therefore, in order to give an exact proof for the material adequacy of the truth definition, one would need to set the discussion at a higher level, because one would need to start talking about the metalanguage, with the use of a meta-metalanguage, and there is no reason to believe that this requirement for a transition to higher order languages would ever stop. This fact leads to the following idea: the material adequacy criterion and the formal adequacy criterion move on almost parallel levels. It can be seen that the formally correct definition is also materially adequate, but if we try to prove it, we encounter very strange difficulties. This seems to be the point where Tarski’s reliance on the correspondence theory of truth starts to fluctuate: on the one hand, Aristotle’s sentence (\textit{To say of what is that it is…}) still sounds enticingly reasonable. On the other hand, the actual formally correct defined notion of truth looks very reasonable as well; however, the system can't include both of them quite easily. At this point one starts to wonder: couldn't there be done a slight modification to the T-sentence, so that it fits better within its formal realm? At which point, the disquotational theory of truth answers: yes, it could.

It is without doubt that Tarski’s conception of truth provides a good ground for the disquotational theories. However, as mentioned before, Tarki’s conception cannot really count as disquotational, and this is especially due to the fact that his actual definition of truth doesn’t have much in common with the disquotational truth schemata. Tarski’s truth predicate is given by an explicit definition, which requires the addition of a metalanguage – being available, therefore, only for drastically restricted fragments of one’s language. Additionally, Tarski’s predicate has a recursive nature, is in need of a satisfaction back-up (instead of backing-up the notion of satisfaction) and, as Field notices\textsuperscript{15}, has a compositional nature. The latter isn’t really a sign of non-deflationism, but it is nevertheless a sign of unaccomplished disquotationism, considering that sentences like \textquotedblleft(p or q) is true if and only if p is true or q is true\textquotedblright don’t follow from the generalised truth schema, but from the recursive definition.

\textsuperscript{14} Alfred Tarski, „The Concept of Truth in Formalized Languages”, in Alfred Tarski, 

Disquotationally-flawed as it may be, Tarski’s truth theory is still the work that had set in motion what was to become the present debates about the status of the T-sentences. One thing is clear: the disquotational sentences have to bear some type of necessity so that they can be disquotational. If the stuff on the right side of the T sentence (the disquotational expression) is not more than merely formally equivalent with the stuff on its right side (the result of the disquotational „process”), then we haven’t really dis-quoted the quotation marks name.

We have several choices for the necessity required by a T-sentences, among which aprioricity and analyticity. There is however a very strong reason not to get too enthusiastic about a prioricity: it is very difficult to find „any satisfactory definition of a priori that is not in terms of another modality”\(^\text{16}\). Therefore, the disquotation schema has to be seen as analytical\(^\text{17}\). Field is talking about analyticity in terms of cognitive equivalence, as the relation that holds between the left side and the right side of an analytic equivalence. And since a sentence’s being disquotationally true must be cognitively equivalent to the sentence itself (as the purely disquotationally use of the truth predicate requires), the T-sentences, in their deflationary acceptation, are bound to be analytic.

Unfortunately, as Halbach notices, Field’s „cognitive equivalence” brings forth a few noticeable problems. The most disturbing ones are the problems induced by Field’s remark that „being cognitively equivalent to” is said not about two sentences, but about two sentence-readings. Assuming that his note: „readings as here defined are identified in computational rather than semantic terms”\(^\text{18}\) should eliminate the problem of characterising a truth predicate for private language, there still remains a question about the compositionality of these sentence-readings. In Field’s sense, the T-sentences tell us that the sentence-reading „«Snow is white» is true” is cognitively equivalent to the sentence-reading „Snow is white”; but if we are interested in the disquotational truth predicator, how can we extract it from the whole of the „«Snow is white» is true” sentence-reading? And assuming that we can do that, how can we be sure that the T-sentence (in Field’s meaning) will tell us something about the truth predicate, and not something about a truth predicate-reading?

Apart from these troubles, Field’s cognitive truth equivalence comes as a modalization of the disquotational sentences, for languages that contain modal

\(^{16}\) Volker Halbach, \textit{op. cit.}, p. 117


\(^{18}\) Hartry Field, „Disquotational Truth and Factually Defective Discourse”, ed. cit., p. 222, fn 1
operators. „But – adds Field – even when the language doesn’t have a modal operator, the left hand side of [the T-sentence] is to be understood as cognitively equivalent to the right hand side” (p. 123). More clearly, Field claims that „each instance of the disquotation schema […] holds of conceptual necessity, that is, by virtue of the cognitive equivalence of the left and the right hand side“ 19. Therefore in his view the disquotation sentences are analytical, and this in the precise sense of conceptual equivalence. Taking this as granted, Field moves forward to the conceptual necessity of the general disquotation principle, and at this point there are two ways to construct the generalisation: by means of substitutional quantification or by using a weak fragment of a substitutional quantifier (with schematic letters as variables and two more rules of inference for them). Both methods use the background supposition that every instance is conceptually necessary.

In Volker Halbach’s habilitation paper *Semantics and Deflationism*, the story begins almost in the same way, with the same premises – that the instances of the disquotation schema are analytically true and can be assumed as axioms; however this not because of the properties of the concept of „analyticity“, but rather because of the disquotational properties of the truth. As such, the infinite instances of the T-sentence can be replaced by the principle saying that the disquotational sentences are truth-analytical – which is to be understood as analytic in the truth predicate 20.

**Questioning the Analyticity: the Lewy Argument**

As we have seen, that which renders to the T-sentences their axiomatic status is the intuition according to which they are warranted. When I say

(1) „It’s raining heavily” is true if and only if it’s raining heavily.

I don’t insert any surplus of content, either on the left hand, or on the right hand side of the equivalence. The two sentences say the same thing. The equivalence (P) is true in every possible world 21. This is what should happen if the equivalence were really, in some way, analytic. But is it really so? Is there no way to imagine a possible world (with a decent degree of possibility) where the sentence might be false? Since „It is raining heavily“ is a sentence that contains indexical elements, it would be tempting to try to find possible worlds that ‘falsify’ our T-sentence. And

---

19 Hartry Field, „*Deflationist Views of Meaning and Content*“, ed. cit., p. 114
20 Volker Halbach, *op. cit.*, pp. 34-40 and 113-134
21 In order to avoid using the incriminated notion, one could also say that the equivalence holds in every possible world.
if one of them turns out to be contingent, the whole principle according to which the T-sentences are analytically true fails, we are forced to decide that the original intuition must have been generously furnished to us by the Deceiving Demon, and the concept of truth has to change its status again.

First Problem

One could say that for a counter-example we don’t even need to apply the truth-predicate to an indexical sentence. Arguments in the style of Casimir Lewy’s seem to provide reasons for believing that no T-sentence is analytic. Let us take the T-sentence built over „Snow is white” as an example. It is possible to imagine a world where the word „white” does exist (like all the other words in the T-sentence), but it means something else, for instance pink. Then our T-sentence would be false in this world. False in one possible world means contingent.

First Solution

The Lewy problem seems to be a form of indexicalisation of the whole language. Since every word might have „meant” something significantly different – or might have been used in a different way, or might be connected to a different stimulus meaning in some world, then every word seems to be an indexical over possible languages, like „I” or „now” are indexicals inside a language. Therefore, once the disquotationalist has solved the problem of local indexicality, a solution for Lewy’s argument will stay as well at his disposal. The current solution for the problem is to make explicit an amendment or refinement implicitly hidden behind the truth predicate: a sentence is not absolutely true if and only if..., but true-as-I-understand-it if and only if... One of the most straightforward examples to illustrate how this works begins with a dialogue between X and Y, seen from the perspective (and understanding) from a third party, Z. X says towards Y: „You are Y”. The T-sentence we are allowed to build while addressing to X is clearly not

(2) „You are Y” is true if and only if you are Y.

but

(3) „You are Y” is true if and only if he is Y.

We are justified to reject the equivalence (2) as a bogus T-sentence because, while uttering it, we had ignored that the truth predicate was supposed to

22 Casimir Lewy, „Truth and Significance”, Analysis, vol. 8, 1947, pp. 24-27
23 This is the solution chosen by Hartry Field and Volker Halbach.
be relativised to our understanding. About the same kind of argument can be brought forward against Lewy’s objection.

Second Problem

In the meantime, considering the deflationary perspective we are situated in, we cannot afford a blind acceptance of the recently introduced concept. Otherwise we should be dealing with a dyadic predicate, „understood as”, that would have to correlate the sentence within quotation marks (or named by the naming function of choice) to an utterly unclear entity, namely my „understanding of” it, while taking into consideration the whole set of indexicals connected to my act of understanding. Besides, there would be the danger of splitting the truth-predicate into the set of {„true as X understands it”, „true as Y understands it”,…}, and that phenomenon would devastate a disquotational theory of truth.

Second Solution

The disquotationalism must therefore give some kind of account of „being true-as-I-understand-it”, and a satisfactory account of it could be furnished by the notion of translation. This is an old idea, indirectly inherited from Tarski’s classical „The Concept of Truth in Formalized Languages”, and considered seriously in Hannes Leitgeb’s study of „truth as translation”.24 Initially the problem of translation had to be taken into attention due to the distinction object language / metalanguage. If we adopt a Tarski hierarchy, then in order to express the T-sentences we need to translate the quoted (or named) sentence from object language into metalanguage, and to mark the equivalence as holding between calling a sentence from the object-language „true” and its translation in the metalanguage. This strategy could also be employed for expressing T-sentences about foreign languages, whereby any sentence containing indexical elements is to be treated as foreign. And then the strategy can be applied universally, since for the trivial cases, or for systems that reject Tarski’s hierarchies, one can use a homophonic variant of the translation function.

Third Problem

The translation seems to be a much better tool than „understanding as”. But even at this point one can still wonder, first, what justification there is to assert the connection between the quoted and the disquoted sentences. This is not a very acute problem, as we could assume, together with Leitgeb, that the concept of

Crăiţă FLORESCU

„stimulus meaning” furnishes a good foundation for translation. A second and quite painful trouble, however, is raised precisely by this unavoidable affinity between truth and translation: how should one deal with the latter, so that one doesn’t meet Quine on the way? And if this can’t be done, what can we do so that the truth predicate doesn’t inherit the indeterminability from the translation? If the translation is indeterminate, how can we choose between

(4) „Snow is white” is true iff snow is white.
(5) „Snow is white” is true iff the colour of snow is white.

Since the right hand side of the equivalences is the translation in my language of the sentence within quotation marks, that is, the sentence-as-I-understand-it, it is unclear why (5) relies on the meaning of „colour” whereas (4) doesn’t have to rely on the meaning of „snow”, so there doesn’t seem to be any way to chose the right instance of the T-sentence between the two variants of denominalisation. Besides, the original idea was that the disquotational theory of truth should provide all there is to know about the truth predicate. If the two equivalences are not significantly distinguishable, then there might be some facts about the truth predicate that the disquotational theory of truth won’t be able to account for.

And thirdly, how can we still get rid in an acceptable way of the distinction language / metalanguage, since „true” is „true-as-I-understand-it”, and that is „true-as-translated-in-my-language”, therefore the T-sentences is expressed completely in my language, while the quoted sentence is not just a sentence in a foreign language I am talking about, but a sentence in a foreign language not conceived as an extension of my own?

The solution to be accepted in order to preserve the intuitive analyticity of the disquotational theory of truth is to combine the theory of truth with a reasonable theory of translation that should allow the deflationist to talk about different sorts of languages. However, it is still hard to see how the theory of truth can avoid the problem of indeterminability, inherited from the very translation it is compelled to use.

Philosophy after Deflationism?

In the end, after all this talk about axioms, analytical equivalences and devices for disquotation, one could still feel slightly disoriented: but wasn’t „truth” that famous notion talking about World, Reality, Existence, Right and many other issues? Wasn’t „truth” one of the fundamental ingredients in epistemological, ontological and many other sorts of philosophical theories? To this question, a deflationist would answer: there must be a mistake, „truth” is not that famous notion.
The disquotational perspective of truth

The notions that answer for World, Reality or Existence are „world”, „reality” and „existence”. „Truth” is only our device for denominalisation, or disquotation. Thus, we need a theory of truth as much as we need a theory that uses names for its own sentences, and if we need a theory of truth, then we will be searching for a theory of denominalisation. On the whole, a deflationary conception of truth does not, nor does it intend to cut out the major philosophical problems, such as that of „Reality” and its concordance with the information provided by our senses, or expressed in our sentences. Issues such as this one don’t have to cease being central to the philosophical thought; they are just no longer connected to the theory of truth.

Bibliography

Volker Halbach, Semantics and Deflationism, unpublished Habilitationsschrift, 2001
Hannes Leitgeb, Truth as Translation, Forschungsbericht der DFG – Forschergruppe Logik in der Philosophie, no. 44, Konstanz, 1999
Casimir Lewy, „Truth and Significance”, Analysis, vol. 8, 1947, pp. 24-27

Crăiţa FLORESCU: assistant researcher and PhD student, „Alexandru Ioan Cuza” University of Iasi. Title of the PhD thesis: The Problem of Interpretation from the Point of View of Formal Logic.