Prolegomenon to Any Consistent Semantics for the Liar

Casper Storm Hansen

Draft of November 8, 2019

Abstract: This paper argues that, given certain mainstream assumptions, the solution to the liar paradox is straightforward, once a few simple distinctions have been made, namely between states of affairs and truths, between the truth property and the truth predicate, and between truth conditions and so-called naïve truth conditions. The solution is really what Chihara would call a "diagnosis", for no claim is made about whether the liar sentence has the truth property or not. That is because the question of whether gap theory, revision theory, or dialetheism (to name just a few options) is correct is an empirical question inessential to the philosophical solution to the paradox.

1 The obvious solution to the liar paradox

The liar paradox arises from the sentence "the liar is not true", when that sentence is called "the liar". If it is true, it seems to follow that it is not true, and if it is not true, it seems to follow that it is true. Since it must be either true or not true, it seems to follow that it is both true and not true.

One possible reaction to this argument is to accept the conclusion that the liar is both true and not true. That is, in effect, to accept that the liar paradox constitutes a philosophical challenge of such an extreme character that it can be used to justify what would otherwise be considered an outrageous view. The principle that is discarded to accommodate the liar is the principle that any state of affairs either obtains or fails to obtain, and not both: with the state of affairs of the liar being true considered to be a counterexample. Rejection of this principle constitutes a revision of standard views that could hardly be more fundamental.

An argument could be made that any coherent alternative account of the liar should be preferred, if it required less extensive revisions to a mainstream web of beliefs. But as it happens, I do not think it necessary to accept such a low standard for alternative accounts. For, as I will argue, there is an obviously correct alternative account—not obvious in the sense that anyone could easily figure it out for themselves, but obvious in the sense that when you see it, you should not be in doubt of its validity. I realize that this seems like an absurd and arrogant claim, given how many previous attempts to solve the paradox there have been. However, most such solution attempts have been formulated under the influence of a certain dogma about what a solution is supposed to accomplish. I will return to what that dogma is

in section 3; but if we free ourselves from it, a simple solution in line with standard background views is readily available.

Another possible reaction to the liar argument is to deny the last premise: that the liar is either true or not true. However, that would imply rejection of the same fundamental principle as before, just with a *gap* instead of a *glut*, which is no less outrageous. By a "state of affairs" I mean a way the world can be (independently of whether and how that way can be described in language). So the principle that any state of affairs either obtains or not— never both, never neither—is absurd to deny.[1] Let us take it for granted from here, and see if we can solve the paradox without calling it into question.

The liar being true is a state of affairs. So it either obtains, or it does not. This is the fact that seems to lead to the contradiction; but by distinguishing between the truth property and the truth predicate, we can avoid that contradiction. Let me first discuss the truth property. As a property of sentences it is obvious what this property consists in: a given sentence has it if and only if (1) truth conditions have been assigned to the sentence by convention and (2) those conditions are satisfied. While it is possible to disagree with this analysis, you would have to be pretty far out of the mainstream to do so. Indeed, it is so minimalistic that almost anyone should be able to accept it. I suppose you are more likely to object that the analysis is superficial and that there is much more to be said about (1) and (2). I will say a bit more in section 2, but it is not strictly necessary, for the minimalistic analysis is sufficient to solve the paradox.

If (1) and (2) are the case for the liar, then it has the truth property or, in other words, the state of affairs of the liar being true obtains. If at least one of them fails, then not.

While truth conditions are assigned to sentences by convention, they are typically not assigned individually, but rather collectively through rules of compositionality. That is, semantic conventions are typically about how words contribute to sentences' truth conditions, and thus only indirectly about the truth conditions themselves. And the semantic convention about how a given word contributes to truth conditions is typically uniform, in the sense that there is a (relatively) simple rule for how that word contributes to truth conditions across sentences. But obviously, that is not so by necessity: we *could* decide, collectively, that from this day forward, the truth condition of "the meat loaf is not in the oven" is that the Moon is full. Exceptions to the simple, compositional rules are—although generally unwanted—possible. Again, if you deny this, you have some very deviant views.

Next, let us look at the truth predicate. It is obviously a predicate that has

---

[1]One might think that whether a state of affairs obtains may be vague, so that there are more than the two options I mention. I happen to believe otherwise, but I do not need that assumption. For the sake of argument, I can allow that a state of affairs obtains to degree 0.5 and fails to obtain to degree 0.5, so to speak. If you believe that this is possible, just replace the liar with the definite liar in the rest of this paper, where "the definite liar" refers to "the definite liar is not definitely true". What I meant to rule out is that a state of affairs obtains to degree 1 and fails to obtain to degree 1, or that the sum in some other way is different from 1.

been added to language with the intention of facilitating description of (and enquiry about, and speculation about, etc.) states of affairs consisting of a sentence having the truth property. That is, the idea was that a sentence of the form "ϕ is true" should have the truth property if and only if ϕ has the truth property. That was the simple, compositional rule intended by the language community. While it may require a considerable amount of work to make that claim entirely precise, it should be obvious that there is a reasonable sense in which it is correct.

In the same sense, the language community intended for the negation to contribute to truth conditions by the simple compositional rule that a sentence of the form $\neg\phi$ has the truth property if and only if ϕ does not have the truth property.

Thus, it was indirectly intended that "the liar is not true" should have the truth property if and only if "the liar is not true" does not have the truth property. By this being the intention *indirectly*, I mean that it would have to be satisfied for the two rules that were *directly* intended to hold without exception to both do so.

However, it cannot be the case that "the liar is not true" has the truth property if and only if "the liar is not true" does not have the truth property, because the liar being true is a state of affairs that either obtains or not. The flip-side of the fact that it is within our powers to make exceptions to general compositional rules is that it is not within our powers to make general compositional rules exception-free, if doing so would conflict with the basic metaphysical law that any state of affairs either obtains or not. If the language community had attempted—say before the general truth predicate and the general negation had been introduced—to make the sentence "aghafyi", with no meaningful proper parts, true if and only if "aghafyi" is not true, then they would have failed. And the language community also failed when it, indirectly and without realizing it, intended for the liar, which *does* have meaningful proper parts, to be true if and only if the liar is not true.

This is quite obvious. Of course, admitting it implies admitting that we humans are imperfect when it comes to designing a linguistic system: our intentions are not always realized. But it is hardly news that we are imperfect in general, so it should not be so difficult to accept that we are also imperfect in this specific regard.

Acknowledging this specific type of imperfection means that we need to make a distinction between the actual truth conditions of sentences and their "intended" truth conditions: that is, the truth conditions they would have if the general compositional rules were exception-free. I will call the latter "naïve truth conditions". The (actual) truth conditions of a sentence are the conditions such that necessarily, given the actual language conventions, the sentence has the truth property if and only if those conditions are satisfied.

Typically, of course, truth conditions and naïve truth conditions coincide. It is equally obvious that they do not and cannot, in the case of the liar sentence. For if they did coincide for the liar, then the liar would be true iff the liar were not true.

That the liar's truth conditions and its naïve truth conditions are different is, I submit, the solution to the liar paradox. With that realization, we see that the liar being true is not, despite appearances to the contrary, a state of affairs that both obtains and fails to obtain. That is all that is required to solve a paradox.

2   Conventional truth conditions

We can add some details and make the solution more concrete by plugging in an analysis of the nature of conventions. I will rely on the analysis by Lewis (1969).

Lewis understands conventions as solutions to coordination problems. An example of a simple coordination problem is that two people need to meet each other once a day, but it does not matter where. The first person will want to go where the second person goes, and the second person will want to go to where the first person goes. If they succeed in finding a place to meet regularly, and meet there because they expect the other person to show up there, they have instituted a convention. In general, Lewis analyzes conventions as arbitrary but self-perpetuating solutions to coordination problems: arbitrary in the sense that there are multiple solutions that, to all the involved agents, are roughly equally beneficial; and self-perpetuating because the choice of one of these arbitrary solutions at an earlier occurrence of the problem will tend to make agents choose the same solution in later occurrences.

In addition to characterizing conventions, Lewis explains how they can be initiated. Returning to the example, the two people do not have to arrive at their convention through explicit agreement; maybe they are unable to do so. They may meet where they do because they both take a chance showing up there, perhaps because they both expect that the other person is most likely to show up there, perhaps because they both expect the other person to consider it most likely that he himself will show up there, etc., through some finite number of higher-order beliefs. The meeting-place may not be intrinsically better than the alternatives; it may simply be that this place has some salient feature that induces the expectations, the expectations about expectations, and so on. Having met there once, the place becomes more salient for the purpose of subsequent meetings, and thus the different orders of expectations become stronger. A convention is in place.

A central insight of Lewis's is that his analysis extends to languages: they are conventional and can arise in basically the same way. Agreeing on the use of a language in a community is a solution to a coordination problem regarding communication. In the case of an indicative sentence, the convention consists of an agreement among a majority of the language users on the conditions under which it is appropriate to assert the sentence. Once we have a convention in place for an indicative sentence, I can, by uttering the sentence, inform you that the world is such that those conditions are satisfied. Thus, according to Lewis, for a sentence to be true is simply for the actual world to be among those possible worlds for which it has been

conventionally agreed (in the actual world) that it is appropriate to utter the sentence in question—appropriate in an idealized sense where we disregard matters of relevance (the sentence may be true but of no interest to the conversation partners), the possibility that the agent has incorrect beliefs, the possibility that in the given situation it is morally obligatory to lie (e.g., to mislead a murderer), etc.

A language community has the power to institute conventions and to decide in which circumstances a sentence is true, i.e., conventionally appropriately assertible; but that does not give the community the power to make a state of affairs both obtain and not obtain. If they try to, they fail. If, for instance, they explicitly agree that "it is appropriate to assert of an object that it is *kratosk* if it is more than 1 metre high, and it is inappropriate to assert of an object that it is *kratosk* if it is less than 2 metres high", then they have not made the sentence "x is kratosk" both true and not true in cases where x is 1.5 metres high. What is most likely to happen in that situation is that their "agreement" results in confusion and fails to govern their linguistic behavior, meaning they have failed to create a convention for the appropriateness of asserting "x is kratosk" in those circumstances. It is also possible that they will all go on to act upon, say, the first half of the agreement, so that the agreement, although not followed, nevertheless causes a convention to be instituted. And there are many other possibilities. But it is not possible that their agreeing has the effect that they all act in compliance with a contradictory convention when confronted with an object 1.5 metres high.

Similarly, a language community does not have the power to make a state of affairs both obtain and not obtain by assigning conventional truth conditions to "The liar is not true". In no language community can that sentence be both conventionally appropriately assertible and not conventionally appropriately assertible at the same time. Thus, in our language community it is not the case that the liar both has and fails to have the truth property.

The solution to the paradox becomes more concrete and tangible when Lewis's analysis is applied, but the solution is not dependent on this analysis being entirely correct. For instance, many dissenters have put forward examples that show that Lewis's definition of convention is either too broad or too narrow—see, e.g., Burge (1975) and Miller (2001)—and these examples do not affect the solution to the paradox. I also do not depend on the concrete claims that are made above about how agents manage to arrive at a particular solution; only that there is *some* mechanism different from explicit agreement that allows them to do so, for example the one explained in Skyrms (1998)—if there were no such mechanism, language would be impossible and there would not even be a liar sentence. And while the primary purpose of language conventions clearly is coordination with respect to communication, they do not have to be solutions to coordination problems in the specific game-theoretic sense that Lewis relies on. Alternatives that could be substituted for Lewis's account include Gilbert's (1989). Only an extreme degree of disagreement with Lewis will lead to incompatibility with the solution. One would have to deny that language is conventional and instead claim that rules of compositionality are backed up by a force so

strong that it can break through the barriers of metaphysical laws.

3 The dogma

I hear an objection: "You have not told us whether the liar has the truth property or not. Any proposed solution to the paradox must give an answer to that question." I beg to differ. That is the dogma I alluded to earlier: that the hunt for a solution to the liar paradox should be conceived of as the hunt for a truth value for the liar sentence. Countless, increasingly complex semantics for formal languages that contain a liar have been proposed with the aim of providing that truth value. But figuring out the truth value is orthogonal to solving the paradox, which consists merely of recognizing that there is a difference between actual truth values and naïve truth values. If you do the former without doing the latter, there are two options (assuming that you respect the constraint that states of affairs either obtain or not). The first is that you conclude that the liar has the truth property, and then you will, in effect, be claiming that the actual truth condition of the liar is satisfied (because otherwise it would not have the truth property) and that the naïve truth condition is not (because the naïve truth condition is that the liar does not have the truth property). However, you will not have made that distinction, so it will seem to your critics that you are contradicting yourself. The second is that you conclude that the liar does not have the truth property: in effect claiming that the actual truth condition of the liar is not satisfied but its naïve truth condition is. And that has the same effect. In fact, this has happened so many times that one could get the impression that semantic theories adhering to the principle that any state of affairs either obtains or not are hopeless. Indeed, it is for this reason that Priest (1987) calls the search for a successful consistent theory a degenerating research program.

Understanding the conceptual difference between naïve truth conditions and truth conditions, and realizing that they must diverge in the case of the liar because of human weakness in the face of metaphysical laws, are the philosophical insights needed to solve the philosophical puzzle that is the liar paradox. *How* they diverge is a contingent matter that depends on the fine details of our language conventions. Whatever the truth value of the liar is—and whatever exact way we have collectively failed in our doomed attempt at making our compositional rules exception-free—it could have been the opposite and we could have failed in a different way. Maybe the truth values *are* opposite in, say, English and Chinese.

To illustrate and exemplify what I have said so far, I will nevertheless comment on three of the most popular theories about the precise semantics of the liar—Kripke's (1975) gap theory, revision theory (Gupta 1982), and dialetheism (Priest 2006b)—not with the aim of declaring which are correct and which are not, but of determining which are *possibly* correct. It turns out that the answer for all three is the same: some versions are possibly correct, and some could not possibly be (given mainstream assumptions).

4   Possible language conventions

I have accounted for the *property* of truth. It is a property that a sentence has if there are conventional conditions for the circumstances under which it is appropriately assertible, and those conditions are satisfied. And about the truth *predicate* I have said that it is a predicate that has been introduced into language with the intention of being able to describe the situation of something having the property of being true.[2] The idea was to have a sentence like "ϕ is true" have the property of being true iff ϕ has the property of being true and to have a sentence like "ϕ is not true" have the property of being true iff ϕ does not have the property of being true. However, intentions are not always fulfilled, and in this case the liar sentence prevents this intention from being fulfilled. The truth conditions for a sentence ϕ are the conditions under which ϕ has the property of being true, and not necessarily the conditions under which a sentence in which the truth predicate is applied to a term referring to ϕ has the property of being true; nor are they necessarily the opposite of the conditions under which a sentence in which the negation of the truth predicate is applied to a term referring to ϕ has the property of being true. The members of a language community can have the intention of making it so, but since that intention is inconsistent (given certain other properties of their language), they cannot succeed.

It is natural to ask what the actual truth conditions for sentences of those forms then are, in full generality. Unfortunately, when an intention is frustrated, there is rarely a way to predict what happens instead merely from knowledge of the intention. If someone intends to build an immovable object and create an unstoppable force, I know that his intention will be not be fulfilled, but there are several possibilities for what will happen instead, namely that he builds an immovable object but fails to create an unstoppable force; that he creates an unstoppable force but not a immovable object; and that he manages neither. Similarly, I do not know what the actual truth conditions of "ϕ is true" and "ϕ is not true" are in all cases. But I can describe several possibilities, and that is what I will do in this section. To argue for gluts in states of affairs from the existence of an intention to create an immovable object and an unstoppable force, one would have to argue that it is more plausible that both are created than that one of the three above-mentioned alternatives ensue. Similarly, to make gluts in states of affairs plausible by appealing to the existence of the liar, one would have to argue that it is more plausible that language users manage to make the liar both true and not true than that language conventions are as in one of the four possibilities described below (or as in one of innumerable other alternatives).

One possibility is that a moderate version of Kripke's (1975) theory is correct: neither "the liar is not true" nor "the liar is true" has the truth prop-

---

[2] *Pace* the deflationists who are so impressed by the fact that the truth predicate can be used to make blind generalisations that they attempt to reduce the semantics of the truth predicate to this utility, and claim that there is no property of truth. See, e.g., Beall (2009, 2), who writes "Our device 'true' [...] was not introduced to name any feature of the world". After a language, initially without a truth predicate, had been introduced into the world, there was a new feature of the world to be named, namely the property of being true.

erty. That is certainly a possibility, for "the liar is not true" having the truth property and "the liar is true" having the truth property are two different states of affairs, and thus might both fail to obtain. The explanation would be, in effect, that the actual truth condition of the liar is as follows: the liar does not have the truth property *and* the liar is grounded. That is one way for the truth condition of the liar to differ from the naïve truth condition in a manner that makes for a possible convention.[3]

Defenders of Kripke may be inclined to say that the truth condition of the liar just is that the liar does not have the truth property, and that the further condition that the liar is grounded is not part of the truth condition but rather a *presupposition* for the sentence to have a truth value, and that that is something else. They may be inclined to do so because phrasing it like that makes it seem like the theory has the virtue of respecting our linguistic intuitions: the truth conditions are exactly what they seem to be. (Similar points apply to the other possibly correct theories about the contingent semantics of the liar discussed below.) However, when phrasing it like that, the defender of Kripke is simply using "truth conditions" differently from how I use it. I defined the truth conditions of a sentence to be the conditions such that necessarily (in those possible worlds that share the actual language conventions) the sentence has the truth property if and only if those conditions are satisfied. Under that definition it is not possible that (1) a sentence ϕ has truth condition c, (2) c is satisfied, and (3) ϕ is not true.

There is also an extreme version of Kripke's theory that goes like this. Semantic facts, such as which truth values sentences have, are metaphysically special. They are (by necessity?) determined by a recursive "process" that begins with an incomplete world in which all states of affairs concerning semantics neither obtain nor fail to obtain, and ends with an incomplete world where *some* states of affairs concerning semantics neither obtain nor fail to obtain. This is a way to avoid my conclusion that there are exceptions to compositional rules and that truth conditions and naïve truth conditions diverge. For according to this theory, the relevant instances of the biconditionals that express the compositional rules are not false, for there is no fact of the matter about whether they are true. And the truth condition of the liar really is that the liar is not true; there is just no fact of the matter about whether that condition is satisfied. While this is coherent, it is also preposterous: there can be no justification for bringing obvious metaphysical

---

[3]I wrote above that truth conditions and naïve truth conditions typically coincide. However, Kripke revealed that for a significant portion of the sentences that contain the truth predicate there must be some mismatch, because those sentences can form part of a contingent liar paradox. Consider Kripke's (1975, 691) example consisting of the sentences "most of Nixon's assertions about Watergate are false" and "everything Jones says about Watergate is true". Both of these sentences cannot have their naïve truth conditions as truth conditions, for under unfortunate circumstances that would lead to paradox. Kripke's suggestion is, in effect, that to get the truth conditions of each, one has to add the condition that the naïve truth conditions are grounded to the naïve truth conditions themselves. Note that unlike in the case of the (non-contingent) liar, these truth conditions are in an extensional sense very close to the corresponding naïve truth conditions: in most circumstances, the sentence will be true iff its naïve truth conditions are satisfied.

8

principles into doubt just to avoid the banal admission that human systems of communication are imperfect.[4]

One version of revision theory[5] would be that the liar is true *now*, false *now*, then true again *now*, and so on. This, too, is a *possibly* correct theory (metaphysically at least, if not epistemically). It implies that the actual truth condition of the liar is that the liar was not true two seconds ago, so the principle that any state of affair either obtains or not (at any given time) is upheld. Also, because the liar sentence is in the present tense, its truth value would never match its naïve truth condition.

As far as I know, no one actually believes the above version of revision theory. Instead, a revision theorist might say[6] that the liar sentence, rather than having the semantic value *true* or *not true*, has the entire revision sequence as its semantic value, or something along those lines. I do not consider this to be a possibly correct theory, because I do not know what the social praxis of communication would have to be like for that theory to be correct. (And that is in spite of me being quite liberal about what I would count as a fact of the social praxis of communication. For instance, I would be willing to consider it a fact of the social praxis of communication that we, in some idealized sense, aim at asserting ϕ and not asserting $\neg\phi$ if ϕ is an unknowable truth.)

Dialetheism, like Kripke's theory, has a moderate version: in this case holding merely that "the liar is not true" and "the liar is true" both have the truth property. This is possible (and, like all possibilities, a violation of our intentions with the truth predicate and the negation). Likewise, there is an extreme version according to which "the liar is not true" both has and fails to have the truth property, which like the extreme version of Kripke's theory avoids the consequence that human language is imperfect, but at an outrageous and disproportionate price.[7]

I also want to mention the possibility that we simply haven't attached truth conditions to the liar: that we have not instituted the relevant kind of convention. Then, by (1) above, the liar would not have the truth property, and hence its naïve truth condition would be satisfied. But that would not make it true.

The thesis that the liar does not have truth conditions is often expressed by saying that the liar does not express a proposition. A traditional problem for such a view is to explain why it *seems* like the liar is meaningful and why we can reason from it as an assumption as if it were. My theory provides an answer: the liar does have naïve truth conditions. They make it seem like it has truth conditions, even if it hasn't, and one can reason from them.

Keeping our system of naïve truth conditions fixed, every possibility for what the corresponding actual truth conditions are implies that there is

---

[4]If I interpret them correctly, Kripke himself represents the moderate version and Field (2008) the extreme one.

[5]For revision theory in general, but not this particular version, see Gupta (1982) and Belnap and Gupta (1993).

[6]Based on conversation and personal correspondence.

[7]I read Priest as advocating the extreme version of dialetheism; see in particular Priest (2006a, 51–54).

some mismatch between the truth conditions and naïve truth conditions of some sentences. Because they do not distinguish between these two, many attempted solutions to the paradox are instead presented as claims about the correct *logic* being different from classical logic. Under one interpretation of those claim, that would be a revision to mainstream views almost as startling as the idea that there are states of affairs that neither obtain nor fail to obtain. But the liar provides as little reason for the former revision as for the latter, for it does not constitute a challenge to this much more reasonable view: classical logic is the correct logic of *actual* truth conditions. For instance, if the actual truth condition of one sentence is the negation of the actual truth condition of another sentence, then exactly one of them is true.

However, there is also a sense in which logic is conventional, for the naïve semantics of the logical connectives and quantifiers do not uniquely determine their actual semantics and thus leaves room for further conventions to settle that. Hence, logic qua syntactical calculus is indeed highly sensitive to convention.

5 T-schemas and expressive strength

What are the pragmatic consequences of the different conventions that have been considered above? For the purpose of ease of communication, it would, for each indicative sentence, be beneficial if that sentence's truth conditions were satisfied (i.e., the sentence were true) iff its naïve truth conditions were satisfied. When this biconditional is satisfied, language users can communicate in the way they naïvely assume they can. When the left-to-right direction is satisfied, hearing the sentence in question being asserted gives a naïve listener exactly the information he or she thinks it gives (assuming the utterer is honest and well-informed). And when the right-to-left direction is satisfied, asserting a sentence conveys exactly the information a naïve utterer thinks it conveys.

A language community where one of the theories described by Kripke is adopted as a convention[8] has, in a manner of speaking, erred on the side of making too few sentences true. That is, the left-to-right direction is satisfied for all sentences. The consequence of the failure of the right-to-left direction for some sentences is a certain expressive weakness. For instance, their liar sentence cannot be used to communicate that their liar does not have the property of being true, even though it not being true is its naïve truth condition.

In an attempt to avoid this kind of expressive weakness a dialetheic language convention could be devised that ensures that the right-to-left direction is satisfied for all indicative sentences. But in the presence of a liar, both directions cannot be universally satisfied at the same time, so there would be failures in the opposite direction. This means that naïve language users

---

[8]For present purposes it does not matter which one, except that what I say here does not apply to the supervaluation version that quantifies over only maximally consistent interpretations.

would not be able to trust true sentence to always convey the information they think they convey.

In a community with the revision convention described above the biconditional can fail in both directions. When the liar has the truth property, it fails in the left-to-right direction for that sentence, and when it does not it fails in the other direction. (However, when the revision sequence has entered into a loop,[9] both directions are satisfied for the stably true sentences, and the set of stably true sentences is a superset of the set of sentences that are true according to the Kripkean convention.[10])

The biconditional is a version of the T-schema, and in these examples it fails. Such failures are not violations of some deep logico-metaphysical necessity; they just result in some communication problems.[11,12]

Another version of the T-schema is this: a sentence is true iff its truth conditions are satisfied. It never fails.[13] A third version is the one mentioned earlier according to which a sentence saying that another sentence is true is true iff that other sentence is true. The term "T-schema" may be ambiguous between these versions and that is the reason why I have refrained from using it up until now, even though using it might have made comparisons between this paper and the existing literature easier by introducing the former as arguing for (among other things) the well-known thesis that there may be exceptions to the T-schema.

---

[9]See Gupta (1982, 45).

[10]Gupta's (1982) revision theory can also be used as inspiration for another convention, namely one where a sentence is true iff it is stably true according to that theory, instead of changing truth value on a daily basis. This convention satisfies the left-to-right direction for all sentences, like the Kripkean convention, and of course the set of stably true sentences is still a superset of Kripke's true sentences. (Disclaimer: neither this convention nor the other Gupta-inspired convention should be interpreted as reflecting Gupta's actual view.)

[11]In addition to being a problem for communication between one person and another, the liar may constitute a problem for thought. I am not sure how severe that problem is. One factor that would be relevant to determining its severity is the extent to which thought takes place in a compositional language.

[12]In addition to the characterisation in terms of the biconditional, a different characterisation of *how* the various conventions fail may be illuminating. They all squeeze the set of sentences that have their naïve truth conditions satisfied between two approximations: a subset of that set, and a superset of that set. Given a Kripkean convention, the subset is the set of true sentences, and the superset is the complement of the set of false sentences. Given a dialetheic convention, the subset is the set of true-only sentences, and the superset is the complement of the set of false-only sentences. Given a revision convention, they are the set of stably true sentences and the complement of the set of stably false sentences, respectively. But note that in all cases, the complement is relative to the universe of indicative sentences of English, and not to the more limited universe of sentences that are admitted as meaningful according to the formal language in question. So a Tarskian (1944) convention also fits in, for each level of the hierarchy: the subset is the set of true sentences, and the superset is the complement of the false sentences. (And the recent proposal by Scharp (2013) does too: the subset is the set of descending true sentences, and the superset is the set of ascending true sentences. This becomes most clear in section 8.3.) Because the set of sentences that have their naïve truth conditions satisfied is convention-relative, it is impossible to "squeeze" so much that the subset and the superset meet and coincide with the target set.

[13]To be precise: *sentences* that are instances of this schema may fail to have the truth property, if their truth conditions are misleading. But their naïve truth conditions never fail to obtain.

Let me reiterate that by comparing moderate Kripkean theories, revision theories, moderate dialetheic theories, and the theory that the liar sentence fails to have truth conditions I have not compared contenders for the title of solution to the liar paradox. That is, I have not compared epistemically possible solutions. I have compared metaphysically possible conventions. The solution resides, so to speak, on a higher level: it consists in realizing that certain languages really are possible conventions in spite of their shortcomings, while a language containing sentences that are both true and non-true is not. If considered as a proposed solution with a claim to platonic necessity, the shortcomings that each of these languages has may be seen as a reason to reject that solution—and since they all have such shortcomings, one may reject them all and end up embracing metaphysical gaps and gluts. But when considered as a possible convention, the shortcomings have no such force: there is nothing unusual about there being no perfect convention available.[14]

I have on a couple of occasions encountered the reaction to this conventionalist solution that conventions are irrelevant and that the problem for the liar sentence has nothing to do with the fact that the meaning of the sentence is conventional. I can see why people would react react that way. Conventions are a complex phenomenon with many aspects, and most of those aspects are indeed irrelevant to the solution (while perhaps the most salient aspects of conventions for those who react in that way). For instance, it is not important to the solution that conventions are created by *groups* of people. Nor is it important that conventions govern *recurring* situations. It is not even essential that conventions have alternatives (although I think that helps with making my point more vivid). What is essential is that conventions are created by humans with limited power. Our power to create a convention for the conditions under which sentences are appropriately assertible is not a power that allows us to make the state of affairs of the liar being appropriately assertible both obtain and not obtain. If we try to, we simply fail and end up with an imperfect language. Those who claim that if the liar is not true, it must also be true, fail to explain who or what the enforcer of this *must* is and how he/she/they/it can make it so hard that it can defy a metaphysical law, because mere mortals certainly cannot. It would seem that, lurking in the background, there is an assumption of a platonic-fregean realm of propositions that are so perfect that they can ignore the most basic rules restricting everything else.

Given certain restrictions, every possible convention will result in some communication problems of the kind described.[15] The kind of conventions we

---

[14]If someone were to insist, against my advice, on calling, e.g., the language described by Kripke a "solution" to the paradox, they should characterize my position as one of pluralism with respect to liar paradox solutions.

[15]A similar point is made by Maudlin (2004), who also uses the concept of a sentence being conventionally appropriately assertible (in his parlance it can be "appropriate [in a 'sui generis' sense] to assert" a sentence according to "rules" we "lay down"—see page 95). However, Maudlin discusses different ways of changing the conventions about what is appropriately assertible against the background of a conception of truth that is independent of conventions. Thus, his full theory is based on a distinction without a difference, but I agree with much of what he says about appropriate assertability.

have considered all (1) keep in place the sentences of English and their naïve truth conditions, (2) allow the truth conditions of a sentence to be identical to its naïve truth conditions when possible,[16] and (3) only provide an auxiliary convention for what the truth conditions are to be otherwise. So, should we get rid of one or more of those restrictions, and make more radical changes to our language to overcome these communication problems? From a pragmatic perspective, I think the answer is "no", because I suspect that the liar problem is a necessary side-effect of the compositionality that makes language efficient and easy to learn; and that any solution aimed at eradicating it altogether would involve sacrifices that far outstrip the gains. However, we could take a less pragmatic and more theoretical view of the matter and focus on in-principle expressive strength, disregarding the usefulness of languages to actual human beings. Then, the question to be asked would be whether it is possible to construct a language convention such that *for any possible state of affairs, there is a sentence that is true iff the state of affairs obtains.* Self-reference would certainly be a challenge for any attempt to achieve such a convention, for among the states of affairs that need to be expressible in the language are the states of affairs of the language's various sentences being true. But it is not clear to me that the revenge phenomenon implies the impossibility of this.[17] First, because the $\forall\exists$-structure of the italicised clause above renders the goal more modest than what is typically aimed at by those who have constructed languages that can accommodate the liar. And second, because the existence of a liar sentence is not a necessary feature of sophisticated languages, not even ones that contain sentences about their own semantics.

We could call the problem of finding such a convention the *liar expressibility problem.* It is closely related to the liar paradox, but not identical. I am offering a solution to the latter, but the former I have only diagnosed. Unlike the paradox, the expressibility problem strikes me as extremely difficult.

6   Dialectical situation

I believe I have argued convincingly against those who believe that there can be gaps or gluts of states of affairs. Let me consider an objection that they might make. They might say that I have not so much argued against them as just assumed that they were wrong (early in section 1 no less) and argued from that as a premise. I have begged the question, that is.

I will answer with an analogy. A scientist who gives a plausible account of the phenomenon of light without invoking the existence of a luminiferous aether is justified in rejecting the existence of the aether. If the scientist, in the course of giving her account, begs the question against the aether theorist by assuming that the aether does not exist, that cannot be held against her if she eventually manages to account for all the data. That is

---

[16]With this qualification, concerning the revision convention: after a sufficient—perhaps transfinite!—number of days have passed.

[17]I set aside the option that there might be uncountably many different states of affairs.

because rejection of aether is more parsimonious than accepting it, so the scientist who rejects it has Ockham's razor on her side.

Similarly, rejection of gaps and gluts of states of affairs is the more parsimonious position, so the burden is on those who accept them to show that we cannot "save the phenomena" without. I have demonstrated that the liar phenomenon can easily be saved. We just have to accept that human beings are imperfect. And we knew that already.

The primary point in favour of my opponents is the alleged failure of every theory about the semantic paradoxes that respects the bivalence of states of affairs. The job of pointing out these purported failures individually, for a large number of the proposed bivalent theories, has been undertaken with admirable diligence by Priest. However, Priest's recurring line of attack against most of those theories consists in (explicitly or implicitly) identifying a sentence ϕ such that it follows from the theory that the view-from-nowhere truth conditions of ϕ are satisfied, while ϕ is not true according to the theory; and concluding on that basis that the theory is wrong; see Priest (1984; 1987; 1993; 1995; 2002; 2005; 2006b; 2007; 2010; 2012), Priest and Routley (1989), and Beall and Priest (2007). Having separated satisfaction of view-from-nowhere truth conditions from truth, we are now in a position to recognise this type of argument as fallacious.

7 The view from nowhere

Given that our language is in general compositional, it is not surprising that we would initially assume that the truth condition of the liar is the result of applying the usual compositional rules to that sentence. But as soon as it is realized that this assumption leads to paradox, the assumption ought to be abandoned. It *is* surprising that this is not obvious and that some are unwilling to do it even after reflection. Why do we resist the fact that there have to be exceptions to certain compositional rules so strongly that we perceive a paradox? I think this psychological question can be answered with a bit of inspiration from Thomas Nagel and his book *The View from Nowhere* (1986).

The relevancy of that book to the liar phenomenon can be seen as follows. True indicative sentences are supposed to represent the state of the world. The totality of true sentences of a given language is supposed to be the most comprehensive and objective representation of the world that the language has to offer. If the language contains self-referential sentences, then the objective representation is also supposed to cover the part of the world that is the language itself. And Nagel's book is all about attempts to objectively represent totalities that include the representer and the means of representation.

Let me first give a brief synopsis. But be warned that this summary is a bad substitute for the book itself. A full understanding of the content of this section may therefore require reading the book.

Human beings start out with a highly subjective conception of the world, i.e., a conception that is largely shaped by the time and place the person happens

14

to occupy in the world, the specifically human perceptual apparatus, and the human way of life. However, we have gradually managed to attain a view on the world that is somewhat removed from that initial, limited vantage point. For example, a subjective human description of something in the world will often contain color terms; but we have found out that a more objective description can be given in terms of electromagnetic radiation. That is a description which would also make sense to other intelligent species that are quite different from ours, and it can form part of an explanation of our subjective color experiences.

The limit of the transcendental impulse towards an objective understanding of the world is referred to by Nagel as the *view from nowhere*. It is an understanding of the world completely independent of worldly facts about the subject itself, qua observer, that distort and limit how the subject views the world. The word "view" should be understood in a very wide sense, as Nagel applies the concept of a view from nowhere not only to the most obvious area of philosophy, epistemology, but also to philosophy of mind, free will, and ethics.

It is a common human mistake to assume that we command a view from nowhere when we do not. One example is a naïve realism according to which the world has only the kind of properties we can directly perceive (e.g., colors) and not others (e.g., electromagnetic radiation). A second example is the assumption of free will, in its most naïve form, when it comes to deliberating which of several possible actions to take. We pretend that we can take "action from nowhere". In order not to make the deliberation seem absurd, we have to avoid seeing the I that deliberates and is about to act as being merely a part of the natural order. We need (the illusion of) autonomy, and that involves seeing ourselves as agents who can influence the causal stream of events from a place that is nowhere in it.[18]

The same kind of naïvety seems to me to be responsible for the entrenchment of the semantic paradoxes. Our naïve beliefs about what the truth conditions of the sentences of English are presuppose that we can take a view from nowhere with our language; that we do not have to take into account that the language itself is part of the world to be described by it; that the effects of instituting a language convention with which to describe the world can be kept isolated from the world. Let me try to explain. Language conventions have an "input" and an "output": the input consists of the facts relevant to whether the truth conditions of a given sentence are satisfied, and the output is the fact that the sentence is true or not true (as the case may be). If the output effects of instituting a new language convention were isolated to an observation point outside of the totality of facts, then *any* condition about the state of that totality could be used as truth condition. That is, if our linguistic practice took place in a view-from-nowhere location separate from the world, we could "view" the world to determine whether the condition

---

[18]Priest (2002) similarly discusses the liar and related phenomena in terms of the dual concepts of *closure* and *transcendence*, with which he makes points connected to Nagel's theme of totalities of some given kind of entities versus entities of that kind outside of that totality. However, whereas Nagel recognizes the conflict as genuine, Priest insists that it is no worse than that it can be accommodated by a dialetheia.

was satisfied and "record" that in the truth or non-truth of the sentence in question, without affecting the world.[19] For instance, we could use the criterion that the liar does not have the truth property in the world and, if that criterion were satisfied, record that fact in the truth of "The liar is not true" in the view-from-nowhere location; and doing so would have no effect on the world and hence not undermine itself. But in fact, the output must have an effect in the world that supplies the input, and that makes it impossible for certain conditions to be our actual truth conditions.

The objectivisation impulse is built into the way we as a language community create conventions for naïve truth conditions for a sentence. It is (implicitly) assumed that the community can choose *any* subclass of the class of all possible worlds to be the worlds wherein the sentence is true (i.e., choose any truth condition for any sentence). But in any possible world where a sentence is true, it has to be true. That is, in any possible world where a sentence is true because its truth conditions are satisfied, so that it has to be true to live up to its job of describing the objective facts, it also has to be true in the sense of it being a fact among the totality of objective facts of the world that the sentence is true. Since the output facts about truth values are not transcendent, but have to inhabit the same world as the input facts, they cannot be in conflict with the latter. Thus, the subclass cannot contain any possible worlds in which the sentence itself is not true.

If the language community nevertheless attempts to assign truth conditions in a way that presupposes that they are standing outside the world that they are creating a language to describe, then they may fail. The truth conditions they have attempted to assign may not become the actual truth conditions. This can come as a surprise to them if their attempt at assigning truth conditions is via compositional rules, so they do not think about the truth conditions of each individual sentence. Then what they attempt to assign as truth conditions is what I have called "naïve truth conditions". Inspired by Nagel, we can instead refer to them using the more descriptive name "view-from-nowhere truth conditions". The view-from-nowhere truth conditions of a sentence are the conditions under which that sentence would be true if all facts (including facts about truth values) were given independently of the facts that are the effects of the language conventions that govern the sentence. Let me unpack that a bit. Let a possible world W have a language community that has made an attempt to assign truth conditions to a sentence ϕ. The facts on which the truth value of ϕ depends are facts in W, and the truth value of ϕ is also a fact in W. But imagine that we instead had two copies of W, W_1 and W_2, and that the facts on which the truth value of ϕ depends are facts in W_1, while the resulting truth value of ϕ creates a fact in W_2 (W_1 is the "viewed world" and W_2 is the "point from nowhere"). Then, the truth conditions that the language community

---

[19]A more familiar way to describe this is to say that if, *per impossible*, we were in possession of a language that were a meta-language relative to all languages, including itself, then any condition could be used as the truth condition of any sentence in that language. But whereas this description is in terms that are particular to the semantic paradoxes, the description in terms of the view-from-nowhere metaphor unifies the psychological explanation for our confusion concerning the liar with Nagel's explanations for a wide range of phenomena.

attempted to assign could be applied even if they are self-referential. (For instance, if the liar is not true in W_1, the truth conditions that the English-speaking language community has attempted to assign to the liar could be applied to that fact to make the liar true in W_2.) The view-from-nowhere truth conditions are the truth conditions that would be used in that fictional scenario.

However, as this is a fictional scenario, view-from-nowhere truth conditions can be used directly as truth conditions only in those cases where it does not matter that the viewed world and the world from which it is viewed are identical. For instance, the view-from-nowhere truth conditions of "Grass is green" can be used as the truth conditions of that sentence because the fact of the truth of the sentence does not affect the greenness of grass. In the remaining cases, the truth conditions have to be given in some other way (although they can be defined *in terms of* the view-from-nowhere truth conditions).

So, to sum up how the concept of view-from-nowhere truth conditions can be used to understand the case of the liar: There is an objective fact about whether the liar is conventionally appropriately assertible or not. If it is, the view-from-nowhere truth conditions of the liar are not satisfied, and if it is not, they are satisfied. In the latter case, the liar itself would naïvely seem to be the right tool for someone who knows that the liar is not conventionally appropriately assertible to communicate that fact to someone who doesn't know it. But implicit in that naïve thought is the presupposition that they have a "nowhere place" in which to communicate, because (per the assumption that the liar is not conventionally appropriately assertible) the liar is not available to them to serve that purpose, according to the conventions of their world.

So the answer to the puzzle of why the liar is so entrenched is that, because of our tendency to naïvely assume that we command a view-from-nowhere perspective on the world, we think of language as a detached medium with which we can describe an objective world that exists independently of it—as if the liar in the detached medium could be true to reflect the non-truth of the liar in the unaffected objective world (or the other way around). But everything must be, by definition, within the world; and thus the liar marks one of the limits of the transcendent impulse. This way of thinking has formed our way of instituting language conventions. The way our conventional truth conditions are primarily given is based on a presupposition of ideal objectivity (similar to the presupposition of ideal objectivity that was shown to be mistaken when the observer effect was discovered in physics), and we have either not formed language conventions that take into account that this is an over-idealisation, or we are not fully aware of the conventions we have formed that do so. Since this presupposition is implicit and not a conscious part of human beings' use of language,[20] it has seemed to us that

---

[20]Or, at least, it rarely is. Philosophers working on the liar are sometimes exceptions. That seems, for instance, to be the case in this quote from Soames (1999, 6, my emphasis), who, commenting on Tarski's theory, writes that its most serious problem is "the irresistible urge to violate the hierarchy's restrictions on intelligibility in the very process of setting it up and describing it. We tend to forget this because *we imagine ourselves*

there were, or ought to be, identity between truth conditions and view-from-nowhere truth conditions; and that is why the liar paradox has seemed to be an *insoluble*. The liar has clear view-from-nowhere truth conditions, so it has been assumed to have clear truth conditions. However, every attempt to specify those truth conditions and the resulting truth value(s) of the liar has in some way or other conflicted with our intuitions about how truth conditions work. But those intuitions arise from the false expectation that view-from-nowhere truth conditions can be identical to truth conditions in all cases, and that the same world can serve both as the viewed world and as the viewpoint from nowhere when they do. Realizing this ought to dispel the psychological *pull* that the paradox can exert, even after one has been made aware of the more elementary elements of the solution presented above.

8   Comparison with Chihara's position

This paper has a lot in common with Chihara (1979).[21] Both Chihara and I solve the paradox by arguing that truth must be like the *kratosk* predicate. But there is also an important difference between truth and *kratosk*, which has to be accounted for to make the solution convincing, and which Chihara failed to account for. Concerning truth, there is an insistent inner voice that says "There is a definite way the world is, and therefore a clear separation of the indicative sentences into those that describe that world (or a part or aspect of it) and those that don't; so there is a definite set of sentences that are true; so there *should* be a legitimate predicate with *all* of the properties we tend to attribute to the truth predicate". Something similar is not the case for the kratosk predicate. There is no intuition that the world of physical things is (clearly and exhaustively) separated into those that are kratosk and those that are not that survives after it has been pointed out that the definition of "kratosk" is inconsistent. To account for the difference, we need two things: first, the concept of view-from-nowhere truth conditions, which sentences predicating truth of some entity have, while sentences predicating kratosk-ness of some entity probably (see section 2) do not; and second, the observation that instituting language conventions in order to be able to describe the world changes the world. Then, we can answer the insistent inner voice: yes (setting vagueness and the conceivable correctness of certain kinds of anti-realism aside), there is a definite set of the indicative sentences that describe the world, in the sense of having their view-from-nowhere truth conditions satisfied. But we cannot make exactly those sentences have the truth property on pain of changing the set of sentences that describe the world (in that sense).

Chihara refers to his own position as "the inconsistency view of truth", but rather than a logical inconsistency, there is something more akin to a performative contradiction involved in the paradoxical uses of the truth predicate. It is not that something can be both true and not true according to what is intended with the truth predicate (as given by how the predicate contributes to view-from-nowhere truth conditions), in the way that something

---

taking a position outside the hierarchy from which it can be described."

[21]The same holds of Eklund (2002) and the descriptive parts of Scharp (2013).

can be both kratosk and not kratosk according to what is intended with the kratosk predicate (as given by its two-part definition). Rather, it is that the "propositional content" (as given by the view-from-nowhere truth conditions) of a sentence predicating truth or non-truth of something can conflict with a presupposition of appropriately asserting that sentence: namely, that the language community in question has made it true in the circumstances where it is asserted, similarly to how the propositional content of a sentence saying that I am asleep conflicts with a presupposition of me asserting it, namely that I am awake.

A further criticism of Chihara is that he rejects the T-schema outright, instead of noting that there are different versions of the schema and that one of them is valid.

Finally, I find the slogan "truth is an inconsistent concept" to be strongly misleading. First, the truth property is consistent: every object either has it or not. Second, the naïve application conditions of the truth predicate are consistent: for instance, the naïve truth condition of "the liar is true" is that the liar has the truth property and that condition is either satisfied or not. And third, the actual application conditions of the truth predicate are consistent: a sentence of the form "x is true" either has the truth property or not.

9   Revenge

We saw above that whatever the exact nature of the mismatch between truth conditions and naïve truth conditions is, it leads to some communication problems. That is bad, for sure, but it is a pragmatic problem for people who want to communicate, and should not be confused with the philosophical problem of the liar *paradox*. I claim that the latter is solved, and that part of the solution is to understand the nature of the former problem.

I maintain this claim even though the pragmatic problem of communication may affect our ability to communicate about the subject of the liar itself. One may try to come up with revenge problems for my proposed solution, for instance by formulating versions of the liar that employ the term "naïve truth conditions" or the term "truth property". However, any such attempt would at most show that pragmatic communication problems also affect some uses of those terms. There is no way to produce a revenge liar that can be used in a sound argument in favor of the conclusion that there is a state of affairs that both obtains and fails to obtain.

Am I nevertheless, in spite of the communication problems, so fortunate that every sentence I have asserted in this paper is true according to actual conventions iff its view-from-nowhere truth conditions are satisfied, so that my attempt at communicating with you, the reader, has not been obstructed? I do not know, because I do not have a detailed account of actual conventions. However, if the actual language conventions should be insufficient for communicating my theory, that would not imply that my opponents are correct. Pointing out that there may be such a revenge problem for this paper does not show that the there are states of affairs that are gappy or glutty. For

this solution, revenge is at most a communication restrainer, not a paradox reviver.

References

Beall, J. (2009). *Spandrels of Truth.* Oxford University Press.

Beall, J. and G. Priest (2007). Not so deep inconsistency: a reply to Eklund. *Australasian Journal of Logic 5*, 74–84.

Belnap, N. and A. Gupta (1993). *The Revision Theory of Truth.* The MIT Press.

Burge, T. (1975). On knowledge and convention. *Philosophical Review 84*, 249–255.

Chihara, C. (1979). The semantic paradoxes. *Philosophical Review 88*, 590–618.

Eklund, M. (2002). Inconsistent languages. *Philosophy and Phenomenological Research 64*, 251–275.

Field, H. (2008). *Saving Truth from Paradox.* Oxford University Press.

Gilbert, M. (1989). *On Social Facts.* Routledge.

Gupta, A. (1982). Truth and paradox. *Journal of Philosophical Logic 11*, 1–60.

Kripke, S. (1975). Outline of a theory of truth. *The Journal of Philosophy 72*, 690–716.

Lewis, D. (1969). *Conventions.* Harvard University Press.

Maudlin, T. (2004). *Truth and Paradox: Solving the Riddles.* Oxford University Press.

Miller, S. (2001). *Social Action: A Teleological Account.* Cambridge University Press.

Nagel, T. (1986). *The View from Nowhere.* Oxford University Press.

Priest, G. (1984). Semantic closure. *Studia Logica 43*, 117–129.

Priest, G. (1987). Unstable solutions to the liar paradox. In S. J. Bartlett and P. Suber (Eds.), *Self-Reference: Reflections on Reflexivity.* Martinus Nijhoff.

Priest, G. (1993). Another disguise of the same fundamental problems: Barwise and Etchemendy on the liar. *Australasian Journal of Philosophy 71*, 60–69.

Priest, G. (1995). Gaps and gluts: Reply to Parsons. *Canadian Journal of Philosophy 25*, 57–66.

Priest, G. (2002). *Beyond the Limits of Thought* (second ed.). Oxford University Press.

Priest, G. (2005). Review of "Truth and Paradox" by T. Maudlin. *Journal of Philosophy 102*, 483–486.

Priest, G. (2006a). *Doubt Truth to be a Liar.* Oxford University Press.

Priest, G. (2006b). *In Contradiction* (second ed.). Oxford University Press.

Priest, G. (2007). Revenge, Field, and ZF. In J. Beall (Ed.), *Revenge of the Liar: New Essays on the Paradox*. Oxford University Press.

Priest, G. (2010). Hopes fade for saving truth. *Philosophy 85*, 109–140.

Priest, G. (2012). Read on Bradwardine on the the liar. In C. D. Novaes and O. T. Hjortland (Eds.), *Insolubiles and Consequences: Essays in Honour of Stephen Read*. College Publications.

Priest, G. and R. Routley (1989). The philosophical significance and inevitability of paraconsistency. In G. Priest, R. Routley, and J. Norman (Eds.), *Paraconsistent Logic: Essays on the Inconsistent*. Philosophia Verlag.

Scharp, K. (2013). *Replacing Truth*. Oxford University Press.

Skyrms, B. (1998). Salience and symmetry-breaking in the evolution of convention. *Law and Philosophy 17*, 411–418.

Soames, S. (1999). *Understanding Truth*. Princeton University Press.

Tarski, A. (1944). The semantic conception of truth. *Philosophy and Phenomenological Research 4*, 341–376.