Davidsonian Semantics and Anaphoric Deflationism*

David Löwenstein†

Abstract

Whether or not deflationism is compatible with truth-conditional theories of meaning has often been discussed in very broad terms. This paper only focuses on Davidsonian semantics and Brandom's anaphoric deflationism and defends the claim that these are perfectly compatible. Critics of this view have voiced several objections, the most prominent of which claims that it involves an unacceptable form of circularity. The paper discusses how this general objection applies to the case of anaphoric deflationism and Davidsonian semantics and evaluates different ways of responding to it (Williams 1999, Horisk 2008 and Lance 1997). Then, three further objections to the compatibility of these theories are assessed and eventually dismissed (Horisk 2007, Patterson 2005 and Collins 2002). It is shown how these considerations shed light on core issues of the debate.

KEYWORDS: Robert Brandom, Donald Davidson, Michael Williams, anaphora, circularity, deflationism, meaning, prosentence, prosententialism, semantics, truth, truth-conditional theory of meaning.

Contents

1 Introduction

It is widely held that the concept of truth plays an indispensible explanatory role in theories of meaning. Deflationists, by contrast, argue that this concept is merely an expressive device, lacking any explanatory power. On these grounds, it has become commonplace to say that deflationism and truth-conditional approaches to meaning are mutually incompatible.

This topic has often been discussed in very broad terms, encompassing many varieties of deflationism and many truth-conditional theories of meaning. The scope of this paper is comparatively modest. I will only focus on the compatibility of one truth-conditional approach to meaning, Donald Davidson's, with one species of deflationism, Robert Brandom's prosententialism or anaphoric deflationism. I will argue that the widespread view that these theories are incompatible – a view shared by Davidson and Brandom themselves – is an illusion fueled by mutual misconceptions of these theories.

To this end, I will start by spelling out what I take anaphoric deflationism and Davidsonian semantics to consist in (section 2). Second, I will try to establish their compatibility by holding Davidsonian semantics fixed and showing how anaphoric deflationism can account for the role of the concept of truth in this approach (section 3). While this claim has been supported, among others, by Michael Williams (1999) and Max Kölbel (2001), its critics have voiced several serious objections. In section 4, I will address the most prominent among these concerns, which is also shared by Brandom and Davidson, namely the idea that combining deflationism with truth-conditional approaches to meaning involves an unacceptable form of circularity. After discussing how this argument applies to the present case, I will comment on different responses to it – by Michael Williams (1999), Claire Horisk (2008) and Mark Lance (1997). Finally, in section 5, I will assess and eventually dismiss three further objections – by Claire Horisk (2007), Douglas Patterson (2005), and John Collins (2002). These considerations will shed light on core issues of the debate and clarify both Davidsonian semantics and anaphoric deflationism, as well as the way they are compatible.

2   Preliminaries

In this preliminary section, I will briefly characterize both anaphoric deflationism and Davidsonian semantics. But I will not discuss *independent*

problems about either of these.[1] My concern is, assuming both theories are prima facie admissible, whether or not they are compatible.

2.1 Anaphoric Deflationism

While there is some disagreement over how best to define deflationist conceptions of truth,[2] I take it that the most fruitful proposal relies on distinguishing two questions (compare Armour-Garb & Beall 2005b).

The substantial question What is truth?

The functional question What is the conceptual role of truth and the linguistic role of 'true' and related expressions?

Traditionally, philosophers take both questions seriously. They give an account of truth, and then employ this account in order to explain its role in thought and talk. I will call these proposals 'substantial theories of truth' or 'substantivism'. By contrast, deflationists think that asking what truth *is* is at least uninteresting and maybe even deeply misguided. They hold that answering the functional question is the only interesting thing to do. In shrugging their shoulders at the substantial question, deflationists are local metaphysical quietists: As for the alleged *nature* of truth, they do not undertake any ontological commitment whatsoever.

Deflationary theories differ in how to make sense of the functional role of truth. As its name already reveals, anaphoric deflationism[3] relies on anaphoric mechanisms. That is, it relies on relations between linguistic expressions that allow some expression tokenings, anaphoric dependents, to

---

[1]For example, I will leave out independent worries concerning anaphoric deflationism. Compare, for example, Wilson 1990 and Båve 2009, as well as footnotes 5, 6 and 8.

[2]Deflationism about truth is often linked to deflationism about reference. This characterization of deflationism easily translates to this and other cases.

[3]For the original presentation and development of the theory, compare Grover *et al.* 1975, Brandom 1988, 1994 and 2002. Note also that anaphoric deflationism about truth has a sibling in anaphoric deflationism about reference, as developed in Brandom 1984 and 1994. These positions are related, but, strictly speaking, independent.

anaphorically inherit the meanings of other expression tokenings, anaphoric antecedents. Pronouns are paradigmatic examples for anaphoric expressions:

(1)   Even though *a friend of mine* strongly believes in *him*, *she* could never convince me to believe in *God* myself.

The tokening of 'she' is anaphorically dependent on its antecedent 'a friend of mine'. It inherits its meaning by *referring back* to that antecedent. Also, the tokening of 'him' is dependent on its antecedent 'God'. In this case, the it refers *forth* to that antecedent and thereby inherits its meaning.[4]

Anaphoric deflationism treats complex utterances formed with expressions like 'is true' as *prosentences*. Hence, the theory is also called 'prosententialism'. Prosentences are anaphorically dependent sentences that inherit their meanings from other sentence tokenings. Truth talk, the claim goes, does not involve attributions of a property but the employment of a unique anaphoric prosentence-forming device – the *truth operator*.[5]

Consider the following example of ordinary truth talk.

(2)   Sarah's self-description is true.

Understood as a prosentence, (2) inherits its meaning from a claim by Sarah, a sentence tokening picked out by the expression 'Sarah's self-description'. Thus, in a context where Sarah has said 'I am a maverick', (2) means:

(3)   Sarah is a maverick.

---

[4]Linguists call the latter phenomenon 'cataphora' and reserve the term 'anaphora' for 'backwards' cases like the former. Cataphora and anaphora are, in turn, subclasses of endophora. Thus, strictly spreaking, the theory is endophoric rather than anaphoric. I will nevertheless stick to the established term 'anaphoric deflationism'.

[5]Prosententialism holds that truth talk is not essentially tied to the expression 'is true', but to the linguistic function of forming prosentences. Since Brandom understands facts as true sentences (e.g. in Brandom 1994, 327), he is committed to say that 'is a fact' is an instance of the truth operator, too. However, he never explicitly *defines* the truth operator, which would then pick out all and only its instances. This is problematic because he relies heavily on distinguishing the anaphoric truth *operator* from the standard idea of a truth *predicate* (in Brandom 1994, 1997 and 2002). In Löwenstein 2010, I have proposed a definiton of the truth operator which solves this problem.

Other instances of truth talk add a quantificational dimension to the story. For example, (4) is understood as (5) and (6) as (7).

(4)   Some of these provocative remarks are actually true.

(5)   There are some sentence tokenings 't' such that (i) 't' is among these provocative remarks, and (ii) actually, t.

(6)   Everything John just said is true.

(7)   For all sentence tokenings 't', if John just said 't', then t.

Given the occurence of the variables both inside and outside of quotation marks, these quantifiers have to be understood substitutionally.[6]

Prosentences are generic in that any sentence tokening can be the antecedent of a prosentence, given that it is specified uniquely, e.g. by describing or quoting it.[7] Prosentences inherit their meanings from the set of their anaphoric antecedents – a singleton in 'lazy' cases like (2) as opposed to quantificational cases like (4) and (6). And prosentences can occur freestanding or embedded in logically complex sentences, e.g. as antecedents of conditionals. Given these resources, the core prosententialist claim is that *every* use of 'true' can be explained in these terms.[8]

Prosentences enhance the expressive power of a language. For example, they allow for rejecting one from a set of claims without specifying which, like in (8), and for endorsing large or even infinite sets of claims, as in (9).

---

[6]Prosententialism therefore presupposes an apt account of substitutional quantification, including a treatment of sentences involving indexicals, context-sensitivity in general and other challenges. Here, however, I cannot discuss, but only point out this commitment.

[7]Prosentences which specify their antecedents by quotation still inherit their meanings from sentence tokenings rather than from sentence types. However, it is possible to make sense of sentence types in terms of sentence tokenings. In section 3.1, I show how.

[8]This implies that the Liar is a prima facie admissible prosentence. While Grover (1977) has attempted a prosententialist dissolution of the Liar, I cannot discuss this problem here. Further, prosententialism also applies to uses of 'true' which are prima facie non-anaphoric as well as to prosentences with antecedents that go forever unuttered, maybe even necessarily so. Even granting Brandom's use of 'anaphora' where he means 'endophora' (compare footnote 4), it remains to be discussed how these cases can be dealt with.

(8)   At least one of Brandom's claims is false.

(9)   All theorems of first-order logic are true.

Further, prosentences can be used in order to endorse or reject claims without knowing what these come down to. Examples of such 'blind' endorsements, as they have come to be called, involve future antecedents like in (10) as well as utterances one has heard, but not understood (compare section 5.3).

(10)   What the editorial of tomorrow's *Times* will say is true.

Finally, anaphoric deflationism also implies an important scheme:

Truth Scheme 'ϕ' is true just in case ϕ.

Unsurprisingly, the left hand side of this biconditional is interpreted as a prosentence. Thus, if the claim quoted on the left is the one used on the right, they will share their meaning by anaphoric inheritance and thereby validate the biconditional. More generally, the scheme holds whenever the claim quoted on the left is synonymous with the one used on the right. Their identity is just the most trivial way for them to be synonymous.

2.2   Davidsonian Semantics

Let me now turn to Donald Davidson's approach to meaning. While there is a lot to say about the intricacies of Davidsonian semantics, I will content myself with a rough sketch and focus on the present issue: the role of the concept of truth.[9] But some aspects will be spelled out in more detail later.

Davidsonian semantics builds on Tarski's theory of formalized languages (1944) in order to characterize the meanings expressed in some object language L. Such theories should enable us to derive adequate instances of Tarski's famous Convention T.

---

[9]I mainly rely on Davidson's 1967, 1973, 1974, 1982, 1986 and 1991 articles.

Convention T 's' is true if and only if p

By replacing 's' with a sentence quoted from the object language and 'p' with a translation of 's' in the metalanguage, we arrive at instances of this scheme – T-sentences – such as:

(11) 'Hunde bellen' is true if and only if dogs bark.

Tarski used the T-sentences in order to explain the concept of truth, relying on the meanings of the object language sentences they quote. For him, truth in a given object language can be defined as whatever fulfills this function for all T-sentences. Davidson, by contrast, turns Tarski's explanatory strategy upside down: He uses the T-sentences in order to explain the meanings of the object language sentences they quote, relying on the concept of truth. In the T-sentences, he claims, the meanings of the quoted sentences are 'given' on the right hand side, in what is commonly called a truth-condition.

Davidson holds that T-sentences are empirical hypotheses about the meanings of object language sentences. But what counts as evidence for or against a given T-sentence? The answer to this question is Davidson's elaborate theory of radical interpretation – that is, of the interpretation of the linguistic behavior of speakers from scratch. In this theory, the principle of charity plays a crucial role: If the interpretation of linguistic behavior is to be possible at all, the interpreter has to understand the interpretee as rational and reasonable as far as possible. This entails that what the interpretee says must be understood as as logically coherent and as true wherever possible. Thus, interpreters must be able to distinguish between what the interpretee believes or *holds* true and what actually *is* true. Radical interpretation involves a basic distinction between mere belief and actual truth.

The principle of charity is crucial because the interpreter has to determine both the meanings of utterances on the basis of their speaker's beliefs and their beliefs on the basis of the meanings of their utterances. Thus, the very possibility of interpreting and eventually understanding a speaker depends

8

on a way to break into this circle – that is, on charity. The underlying picture of this argument is what Davidson calls 'triangulation': The interpreter understands the interpretee as causally situated in the very world she herself is causally situated in. In canonical situations, she can assume that the interpretee's utterances concern the very objects she perceives herself. This enables her to arrive at first tentative hypotheses, first T-sentences, which can be used as an entering wedge into interpreting the language as a whole.

3   Compatibility

Having set the stage, I will now go on to show why anaphoric deflationism is compatible with Davidsonian semantics. Starting from my rough characterization of Davidsonian semantics, I will go through the points where the concept of truth occurs in this theory – in the T-sentences themselves (sections 3.1 and 3.2) and in the theory of radical interpretation (section 3.3).

3.1   Prosentences in T-sentences

I think that prosententialism makes perfect sense of the role of truth in the T-sentences. For example, the left hand side of (11) is understood as an embedded prosentence which inherits its meaning from an object language utterance of 'Hunde bellen'. The whole T-sentence (11) asserts the equivalence of this prosentence with the sentence used on the right hand side, 'Dogs bark'. Thus, if (11) is true, 'Dogs bark' is equivalent not only with the prosentence, but also with its anaphoric antecedent, the object language utterance of 'Hunde bellen'. This is precisely what T-sentences are supposed to do in Davidsonian semantics. In the words of Michael Williams, a prosentence is an "ideal candidate" for this job (1999, 557).

Really? While Davidsonian semantics is intended as a theory of meaning for sentence *types* rather than individual sentence tokenings, prosentences inherit their meanings from sentence tokenings rather than types. And the

T-sentences are supposed to follow from a finite set of *axioms* on the basis of the *structure* of the object language sentences, while prosentences inherit their meanings from sentence tokenings as wholes. However, both of these problems can be solved within the prosententialist setting.

In order to see how, recall first that prosentences inherit their meanings from the *set* of their antecedents, and consider second that sentence types can be understood as sets of sentence tokenings. Thus, all a prosential reading of the T-sentences requires is an order of explanation that starts with individual sentence tokenings and deals with the question of their possible type-identity later. But this is just the order of explanation radical interpreters have to pursue *anyway* for reasons *independent* from prosententialism. At the outset, when looking at foreign utterances in canonical situations, it is an open question whether several such utterances eventually turn out to be equivalent – whether they turn out to be tokenings of the same type. Interpreters have to stipulate which ones are tokenings of the same type – which utterances are equivalent – and test these stipulations empirically.[10]

This process is already built into the practice of radical interpretation. When I use a T-sentence in order to interpret some utterance, I presuppose that that utterance is an utterance of the sentence quoted in the T-sentence – that the tokening I interpret is a tokening of the sentence type for which the T-sentence gives a meaning. If the T-sentences are understood prosententially, this stipulation can be put as follows: I presuppose that the tokening I interpret is a member of the set of anaphoric antecedents of the prosentence on the left-hand side of the T-sentence. Thus, anaphoric deflationism is compatible with the idea that T-sentences give the meanings of sentence types. They give the meanings of sentence types *in that* they give the meanings of

---

[10]The primacy of tokenings over types in Davidsonian semantics is not limited to the first steps in radical interpretation (compare Davidson 1986). Because we always have to take into account the possibility of irony, metaphor, malapropisms and other phenomena, "the theory we actually use to interpret an utterance is geared to the occasion" (Davidson 1986, 441) – it is geared to sentence tokenings at a given time for a given speaker.

all and only those sentence tokenings which are tokenings of that type.

The same resources can also explain how prosententialism can accomodate the axiomatic character of Davidsonian semantics. For reasons independent of truth talk, radical interpreters start off with T-sentences about object language sentences *as wholes.* Which axioms about the internal structure of object language semtences are appropriate to unify this data is a further question. Interpreters have to stipulate how the object language sentences are structured, stipulate axioms in order to account for known T-sentences and test these axioms by deriving and testing hitherto unknown T-sentences.

These axioms concern the *structure* of sentence types, in particular their phonetic and syntactic structure. Given that sentence types can be understood as sets of sentence tokenings, these axioms concern the structure of those tokenings which are members of the set in question. They specify which structural difference of an utterance makes which difference in meaning. On a prosentential reading of truth talk, such a set of axioms makes it possible to derive theorems including prosentences on the basis of the structure of the sentence tokenings these inherit their meanings from.[11] Thus, prosententialism can account for the axiomatic character of Davidsonian semantics.

3.2 Expression and Explanation

However, Davidsonian semantics needs the T-sentences for a specific job, namely, explaining the meanings of the object language sentences. Davidson himself and the vast majority of his interpreters have taken this claim as follows: The theory explains the meanings of the object language sentences in terms of their truth-conditions. In short, it explains meaning in terms of

---

[11]Of course, since some expressions are context-sensitive, interpreters need a set of axioms which includes contextual differences between utterances. But this does not threaten my point. In these cases, the set of anaphoric antecedents from which a prosentence inherits its meaning is picked out in terms of contextual features alongside structural ones. However, the general problem of context-sensitivity is beyond the scope of this paper.

truth. However, deflationism implies that truth is an expressive rather than a substantive concept, that "the function of truth talk is *wholly* expressive, thus *never* explanatory." (Williams 1999, 547) Hence, it seems like no deflationist account of truth can provide what Davidsonian semantics requires.

Let me take a step back. How should we understand the distinction between expression and explanation in the first place? Apparently, it is at the core of the debate on deflationism and semantics, but still it often remains implicit. In my view, the best way to make sense of this distinction is by analogy to the role of logical vocabulary in explanations like this one:[12]

(12)   An only child is a person without siblings.

I take it that this is an explanation of the concept of an only child in terms of exactly *two* other concepts, the concepts of a person and a sibling. Of course, negation and conjunction also play a crucial, even indispensable role in (12). But their role is just to express the *kind* of explanatory relation between the concept of an only child on the one hand and the concepts of a person and a sibling on the other. (12) does not explain the concept of an only child *in terms of* the concepts of negation and conjunction *and* of a person and a sibling. Rather, it explains that concept in terms of the concepts of a person and a sibling, and it employs the expressive resources of negation and conjunction in order to accomplish this.

I think that cases like this provide us with a sufficiently firm grip on the distinction between explanatory and expressive roles of concepts within explanations. The explanandum is explained *in terms of* all and only those elements which play explanatory roles. And the *kind of explanatory relation* between those elements is expressed by additional concepts which play expressive roles. These concepts express the *form* of the explanans, but they do not provide its *foundation*. They are not *in terms of what* the explanans

---

[12]This analogy ties in with Brandom's broad notion of logical vocabulary, which explicitly includes semantic and intentional vocabulary (Brandom 1994 and 2008). However, my arguments here do not depend on such views on these additional topics.

purports to explain the explanandum, but *how* it purports to do so.[13]

Let me now apply this distinction between explanatory and expressive roles of concepts within explanations to the T-sentences.

Substantivists hold that (11) is an explanation of the meaning of 'Hunde bellen' in terms of the proposition that dogs bark *and* in terms of truth. It explains the meaning of 'Hunde bellen' in terms of the fact that 'Hunde bellen' has the property truth just in case dogs bark. By contrast, the logical concept of bisubjunction only plays an expressive role. (11) is not an explanation of the meaning of 'Hunde bellen' in terms of bisubjunction *as well as* the proposition that dogs bark and truth. Rather, it explains it in terms of the proposition that dogs bark and of truth, and it employs the expressive resources of the concept of bisubjunction in order to accomplish this.

By contrast, deflationists hold that the concept of truth only does expressive work in the T-sentences.[14] Accordingly, (11) is an explanation of the meaning of 'Hunde bellen' in terms of exactly *one* element, the proposition that dogs bark. Granted, the concept of truth plays an indispensable role in (11). But its role is comparable to that of bisubjunction: These concepts merely *express* the explanatory relation between the meaning of 'Hunde bellen' and the proposition that dogs bark. (11) does not explain the meaning of 'Hunde bellen' in terms of the concepts of truth and bisubjunction *as well as* the proposition that dogs bark. Rather, it explains it in terms of the proposition that dogs bark, and it employs the expressive resources of the concepts of truth and of bisubjunction in order to accomplish this.

For everything I have said about Davidsonian semantics, the choice be-

---

[13]Of course, this issue should be discussed much more thoroughly than I can provide in the context of this paper. For example, we should notice that the same explanatory and expressive resources can still lead to different explanations – say, "An only child is a non-person with siblings" instead of (12). Further, even though negation and conjunction play merely expressive roles in explanations like (12), they might very well play explanatory roles *elsewhere*, for example in defining the material conditional. By contrast, a defining feature of deflationism is that the concept of truth can *never* play an explanatory role.

[14]Brandom (2002) grants a role for truth in such 'local explanations of content', but not in 'global explanations of contentfulness' (compare section 4).

tween these two stories is entirely optional. The fact that the T-sentences explain the meanings of the object language sentences they quote leaves it open whether or not truth carries some of the explanatory weight. Thus, anaphoric deflationism is compatible with this idea. The whole explanatory burden can be put on the metalanguage proposition 'p'.

3.3   Truth in Radical Interpretation

Truth also plays important roles in the core of Davidsonian semantics, the theory of radical interpretation. Two of these issues stem from the principle of charity and have already been discussed by Williams (1999, 559ff.). However, the idea of triangulation brings up an important third point.

First, it seems as if truth is involved in explaining logical form and logical consistency. But while logic makes *use* of the concept of truth, it is entirely neutral on the debate about *theories* of truth. To say that 'A' and 'B' are logically inconsistent because, in virtue of their logical form, they cannot both be true, is just to say that we can know that 'not-(A-and-B)' solely in virtue of the logical forms of 'A' and 'B'. We can learn this, for example, by seeing that 'A' implies 'not-B'. Davidson relies on standard first-order logic in order to specify the logical forms of sentences. But nothing about first-order logic implies a commitment to a substantive conception of truth. It does not call for more than the expressive functions of that concept.

Second, what an interpretee says should be interpreted as true whenever possible in order to understand them as rational and reasonable. This is crucial because the assumption of truth in canonical situations provides the much-needed entering wedge into the holistic circle of the interpretee's beliefs, desires and the meanings of her utterances. However, Davidson explicitly says that the standard for this assessment is *not* truth simpliciter, but what the interpreter *believes*, what she *takes* to be true. After all, interpreters can only work with what they have got. In trying to maximize rationality, the principle of charity aims at maximizing truth. But, in a catch-phrase, it aims

14

at maximizing truth *in that* it aims at maximizing agreement.

Finally, however, triangulation brings a notion of truth back in which *transcends* mere agreement. Interpreters do not only have to make the beliefs they attribute to their interpretees agree with their own. They also have to conceive of both these and their own beliefs as capable of failing to be true against an objective standard of assessment. This is a fundamental realist distinction between what is true and what is merely believed. However, it is entirely innocent. Most importantly, it does not amount to a substantivist conception of truth, say, in terms of correspondence. Deflationists happily agree that there is a difference between *merely* believing that dogs bark and *truly* believing it. They simply add that truly believing that dogs bark boils down to believing that dogs bark while dogs actually bark. Again, the work done by the concept of truth is expressive rather than substantive.

I conclude that what Davidsonian semantics needs from a conception of truth can be provided by anaphoric deflationism.

4 Circularity

Ever since Michael Dummett (1959), deflationism and truth-conditional approaches to meaning have widely been pictured as incompatible because combining them would lead to an unacceptable form of circularity. In this section, I reconstruct the general pattern of such arguments and show why they cannot be applied to the present case (section 4.1). Then, I take a brief look at Claire Horisk's (2008) and Mark Lance's (1997) answers to the charge of circularity and argue that my strategy is superior to both (section 4.2).

4.1 Why the Circularity Objection Misfires

What I take to be the most plausible version of this kind of argument can be expressed by the following schematic argument, the Circularity Objection:[15]

1. If deflationism is true, then the concept of truth is explained in terms of the linguistic mechanism M.

2. M can only be explained in terms of meaning.

3. For all concepts, if α is explained in terms of β, and β can only be explained in terms of γ, then γ cannot be explained in terms of α.

---

4. If deflationism is true, then meaning cannot be explained in terms of the concept of truth.

Premise one is a general characterization of deflationism that follows from what I said above (section 2.1): Any variety of deflationism will claim, first, that the only thing to say about the concept of truth is to say how it functions, and second, that it functions according to some linguistic mechanism M. While disquotationalists claim that the functions of truth can be explained in terms of quoting and disquoting sentences, prosententialists hold that they can be explained in terms of anaphoric relations between sentence tokenings.

Premise two is a second general claim about deflationism, but it does not directly follow from a general understanding of this position. However, this case is perfectly clear when it comes to prosententialism: Since the anaphoric relations appealed to are *defined* as relations of inheritance of meaning, explaining anaphoric relations requires appeal to meaning.

Finally, premise three expresses a version of the principle that philosophical explanations of concepts should not be circular. It states that, if we explain, say, justice in terms of fairness and fairness in terms of moral equality,

---

[15]This can be taken as an explication of what Horisk calls "Explanatory Chicken-and-Egg Arguments" (2008, 284ff.). Among many, John Collins (2002) argues like this.

we cannot rely on justice to explain moral equality. Intuitively, these concepts might then stand in a relationship of mutual illumination, but there would not be a full-blown *explanation* anymore.

The present issue, however, concerns anaphoric deflationism and Davidsonian semantics. Let me fill in the first two premises of the schematic Circularity Objection and formulate a concrete argument about this case:

1. If anaphoric deflationism is true, then the concept of truth is explained in terms of anaphora.

2. Anaphora can only be explained in terms of meaning, namely in terms of inheritance of meaning.

3. For all concepts, if α is explained in terms of β, and β can only be explained in terms of γ, then γ cannot be explained in terms of α.

---

4. If anaphoric deflationism is true, then meaning cannot be explained in terms of the concept of truth.

Some proponents of anaphoric deflationism use the conclusion of this argument in order to refute explanations of meaning in terms of truth. They affirm the antecedent and apply modus ponens. Among others, Brandom (1994, 2002) explicitly does so. Proponents of truth-conditional explanations of meaning, by contrast, can deny the consequent and apply modus tollens. One of those who explicitly do so is Donald Davidson (1990). In any case, if the argument is sound, these theories are incompatible.

But, despite its widespread appeal, the Circularity Objection fails because it relies on a misconception of Davidsonian semantics (compare Williams 1999 and Kölbel 2001). The objection establishes an incompatibility between anaphoric deflationism and explanations of meaning in terms of truth. But, given the considerations of section 3, the commonplace characterization that Davidsonian semantics explains meaning in terms of truth is simply mistaken.

The fact that the concept of truth is *used* in the explanation does not entail that this concept does any of the explaining. The uses of 'is true' can all be accounted for in terms of the expressive role of truth-talk. Davidsonian semantics is a truth-conditional approach to meaning, but it does not explain meaning in terms of truth. It explains the meanings of object language sentences in terms of the practice of interpreting them in a metalanguage. The explanatory work is done by the theory of radical interpretation alone. Thus, even if the Circularity Objection is sound, it simply misfires, because its conclusion does not apply to Davidsonian semantics.[16]

4.2 Other Proposals

Thus, the Circularity Objection fails because it falsely assumes that truth plays an explanatory role in Davidsonian semantics. Claire Horisk (2008) and Mark Lance (1997) have also argued against this objection. In this section, I will briefly show why Horisk's argument depends on the argument just given, and why Lance's argument cannot be applied to the present case.

Horisk questions premise three of the Circularity Objection, what she calls the 'Priority Assumption'. She distinguishes between two readings: On a metaphysical reading, circular explanations are unsound because they "attempt to break down concepts into simpler and more fundamental concepts" (Horisk 2008, 291) and because one concept cannot be both more *and* less simple than a second one. On an epistemological reading, a circular expla-

---

[16]There might even be a second argument for this conclusion. Davidsonian semantics explains what object language sentences mean, but does this amount to an explanation of meaning in general? Davidson himself says that his theory 'illuminates' the concept of meaning (1973, 1990, 1996), but it is not clear how this metaphor should be understood. Does it mean that "meaning just is whatever the practice of interpretation reveals" (Williams 1999, 553)? Or is the theory intended to deepen and clarify an independently available concept of meaning (compare Horisk 2008)? This is a complicated issue which I cannot fully account for in the present context. Maybe the best account of the metaphor of the 'illumination' of meaning provides a further reason to reject the idea that meaning is explained in terms of truth. However, given that truth's role has turned out to be expressive rather than explanatory, there is sufficient reason to reject this idea anyway.

nation is unsound because the target concept turns out to be involved in its own understanding, making such an understanding impossible in the first place (compare Horisk 2008, 292). However, Horisk argues that both kinds of explanation do not *have* to take this reductive form and that Davidson's explanation of meaning does not do so. He spells out how meaning is connected to other core concepts such as truth, belief and desire. This does not *reduce* meaning to any of these, but still deepens our grasp of this concept. "[H]e does not offer a reductive analysis, he offers a reciprocal analysis, that enriches our understanding of meaning by outlining the connections between the concepts of truth and meaning, and other concepts." (Horisk 2008, 291)

But why exactly is a reciprocal explanation compatible with deflationism? And why exactly can reciprocal explanations be sound, even if they are circular? Horisk explicitly considers a version of this problem:

> [D]eflationists might worry that the account of truth will not be truly deflationary if they have any truck with a reciprocal analysis. But the worry may be unfounded. The issue must turn on the role of truth in the reciprocal account. If truth's role in the reciprocal analysis is restricted to logical and expressive functions, as Kölbel and Williams argue, then perhaps there is no cause for concern. (Horisk 2008, 291f.)

But if this is true, Horisk's argument against the Circularity Objection depends on the soundness of the argument I gave above. The reason why a reciprocal explanation in terms of truth is sound even if circular is that truth only plays an expressive role. But if truth only plays an expressive role, there is no reason to speak of an explanation *in terms of truth* in the first place. This defeats the Circularity Objection *directly*, without Horisk's detour via reciprocal forms of explanation. I take it that, other things being equal, evading such an unnecessary detour is more elegant and less misleading.

I shall now turn to Mark Lance (1997), who has presented an interesting general assessment of anaphoric deflationism including an argument against

the Circularity Objection. He starts with the observation that anaphoric deflationism validates the truth scheme (compare section 2.1):

Truth Scheme 'ϕ' is true just in case ϕ.

Substantivists can be generally characterized as claiming that there is some substantial property – T, say – which applies to all and only truths. Realists might spell out 'T' as 'correspondence to a fact', antirealists as 'maximal coherence', and so forth. The general substantivist position, however, is:

Substantivism 'ϕ' is true just in case 'ϕ' has the substantial property T.

Lance's point is simply this: Anaphoric deflationism does not contradict Substantivism. Rather, the Truth Scheme *transforms* Substantivism into:

Substantivism* ϕ just in case 'ϕ' has the substantial property T.

Thus, he thinks that there is no incompatibility between Substantivism and anaphoric deflationism. By the same token, if truth was supposed to play any explanatory role in a theory of meaning or elsewhere, proponents of Substantivism can rely on the substantial property T instead. By cutting out the middle man, the concept of truth, all the explanatory work is done by what substantivists claim truth *consists in*, the substantial property T. For example, if I hold that truth consists in correspondence, I can leave out truth and explain meaning directly in terms of correspondence-conditions.

But unfortunately, we cannot apply Lance's argument to Davidsonian semantics. It requires that the truth-conditional theory of meaning is supplemented with a version of Substantivism including a suitable truth-property 'T'. But Davidson never subscribes to such a view and even explicitly rejects it in his latest work on truth (1990, 1996). For him, truth is an "indefinable concept" (1996, 265) and deflationists and substantivists alike are victims of "the same ugly urge to define" it (1996, 276). But if we cannot identify 'what truth consists in', no substantive property 'T', Lance's argument fails. For Davidson, truth cannot be a middle man one might also do without.[17]

---

[17]While Lance's argument fails when it comes to Davidsonian semantics, it may still

20

5 Compatibility Defended

I will now discuss three further objections against the compatibility of David-sonian semantics and anaphoric deflationism.[18] My responses to these worries will shed light on the modal status of the T-sentences (section 5.1) and the distinction between metalanguage and object language (sections 5.2 and 5.3), and establish the unity of the concept of truth accross languages (section 5.3).

5.1 Expressive Inadequacy

Claire Horisk (2007) claims that a deflationary notion of truth cannot meet the expressive task it must be able to accomplish in Davidsonian semantics. Here, I will only discuss the version of this Expressive Inadequacy Argument, as she calls it herself, which concerns anaphoric deflationism (2007, 547ff.). The first step in this argument is to show that Davidsonian semantics requires a specific modal status for the T-sentences. They have to be contingent.

Horisk develops this first step in the context of Max Kölbel's (2001) defense of the compatibility of deflationism and Davidsonian semantics, which also includes the 'biconditional doctrine', the 'first dogma' of Davidsonian semantics. Roughly, this is the idea that T-sentences can '*give*' the meanings of object language sentences by specifying truth conditions, but that we cannot *say* what they mean directly. That is, we cannot use T-sentences such as (11) in order to derive what we might call M-sentences such as:

(13) 'Hunde bellen' means that dogs bark.

Kölbel argues that if the T-sentences meet some independently important requirements, it is possible to infer M-sentences, after all (2001, 618ff.). On this

---

succeed for other truth-conditional theories of meaning. In Löwenstein (2010), I argue that Lance relies on an incomplete understanding of substantivism, but that a more nuanced discussion of the relationship between substantivism and anaphoric deflationism reveals an important sense in which his argument is sound, after all.

[18]I ignore a prominent objection by Bar-on, Horisk & Lycan (2001), the Determination Argument, since Douglas Patterson (2007) has already shown very clearly why it fails.

basis, Horisk can easily show that T-sentences must be contingent: Clearly, M-sentences are contingent. Thus, if we can sometimes correctly derive M-sentences from T-sentences, these T-sentences have to be contingent, too. Otherwise, these derivations would not be correct in the first place.

Later, however, Horisk points out that this step does not depend on Kölbel's point: Whatever supported the claim that M-sentences have to be contingent *directly* supports the claim that T-sentences have to be contingent, too (compare Horisk 2007, 550). But she then leaves it to the reader to figure out how exactly the argument is supposed to generalize. However, I take it that she is correct, and that the argument can be put as follows:

In Davidsonian semantics, T-sentences are empirical hypotheses aimed at making as much sense of the interpretee's linguistic behavior as possible. As empirical hypotheses, they constantly have to be reevaluated in the light of new observation and in the light of other T-sentences. Furthermore, they are used in order to explain the meanings of the object language sentences they quote. But the meanings of sentences are arbitrary, thus contingent. One and the same utterance – whether oral or in writing – might perfectly well mean something different from what it actually means. Thus, as empirical hypotheses, and as hypotheses about contingent matters (meanings), whether or not the T-sentences themselves are true has to be contingent, as well.

Horisk is aware of the fact that the modal status of a language's semantic features is a matter of sophisticated philosophical discussion (2007, 550ff.). While I cannot discuss this topic here in more detail, I maintain the intuitively plausible claim that sentences might have meant something different from what they actually mean.[19] Those who reject it will thereby reject the first step of Horisk's argument. Since I grant it, I will have to show what

---

[19] That the meaning of "No bachelor is married" is contingent entails that it is contingent that "No bachelor is married" is true. After all, it might have meant that cats can fly! But this is compatible with the idea that it is necessary that no bachelor is married. The existence of possible worlds where "No bachelor is married" means that cats can fly is compatible with the fact that no possible world contains a married bachelor.

is wrong with her second step, the idea that anaphoric deflationism violates the constraint that T-sentences are contingent (compare Horisk 2007, 549).

Horisk argues that (11) is not merely a true T-sentence, but that it is necessarily true since the prosentence on its left-hand side inherits the meaning of the quoted sentence 'Hunde bellen'. 'Hunde bellen' means that dogs bark, hence the prosentence means that dogs bark, just like the sentence used on the right-hand side. Thus, both sides of the biconditional have the same meaning. And a claim that uses two synonymous expressions and asserts their equivalence is necessarily true. "Since the left-hand side and the right-hand side [..] have the same content [...], they are guaranteed to have the same truth-value in all possible worlds." (Horisk 2007, 549) Thus, prosententialism violates the constraint that the T-sentences have to be contingent.

This may look like strong argument, but it fails eventually. In order to see why, recall the motivation of the constraint that T-sentences have to be contingent: They have to be contingent *in that* their truth depends on the meanings of the object language sentences they quote. But Horisk is wrong when she claims that anaphoric deflationism cannot account for this fact.

Here is why. Intuitively, (11) is a true T-sentence because 'Hunde bellen' means that dogs bark. This is contingent, since 'Hunde bellen' might have meant that cats can fly. In this case, (11) would be false. Given that this is possible, the truth of (11) is contingent. The prosententialist version of this story is this: Since 'Hunde bellen' means that dogs bark, the prosentence on the left hand side of (11) which inherits this meaning also means that dogs bark. But it is contingent that 'Hunde bellen' means that dogs bark. Thus, it is *equally* contingent that the *prosentence* means that dogs bark. 'Hunde bellen' might have meant that cats can fly. Then, the prosentence would also mean that cats can fly, and (11) would be false. Because it is possible that the prosentence means that cats can fly since it is possible that 'Hunde bellen' means that cats can fly, the truth of (11) is contingent.

Horisk fails to see this because she thinks that "on a view like Brandom's,

the content of the left-hand side tracks the actual semantic features of the structually described sentence to which it refers" (Horisk 2007, 550). On this view, it is *necessary* that "Hunde bellen' is true' means that dogs bark, because its antecedent, in *all* possible worlds, is an utterance of 'Hunde bellen' from the *actual* world. But this misconstrues anaphoric deflationism. The antecedent utterance of 'Hunde bellen' is picked out *structurally* rather than semantically – that is, in terms of which noise or which squibble was produced or perceived where and when. But it makes perfect sense to say that the same utterance exists in other possible worlds and that it can be picked out in the same way,[20] while its meaning varies across possible worlds.

In sum, Horisk's worry is that prosentialism makes true T-sentences necessary because it makes them true in virtue of meaning. But there *are* possible worlds in which the meaning of the left hand side is *not* the same as the meaning of the right hand side. Where 'Hunde bellen' means that cats can fly, (11) is false. In a sense, true T-sentences are *indeed* true in virtue of meaning, but not in virtue of their *own* meaning alone. They are true in virtue of the meaning of the object language sentence they quote. But this is not a problem, but what we wanted in the first place.

5.2   The Metalanguage Objection

Let me now turn to Douglas Patterson's (2005) argument against the compatibility of Davidsonian semantics and deflationism – what I call the Metalanguage Objection.[21] Patterson argues that what is at stake becomes clearer "if we pay attention to the distinction between the object-language *for which* an account of truth and meaning is given, and the metalanguage *in which* this account is given" (2005, 276). In particular, we need to distinguish between accounting for the object language truth-concept and accounting for

---

[20]On identity criteria for utterances (tokenings) in terms of their structures, recall sections 2.1 and 3.1. In principle, these might even be cashed out in physical terms.

[21]John Collins also hints at an argument along these lines (compare 2002, 520f.).

the metalanguage truth-concept (compare Patterson 2005, 279ff.).

According to Patterson, deflationists claim that the meanings of object language sentences and applications of the concept of truth to these sentences are equivalent in what he calls 'basic meaning' – paradigmatically, in free-standing uses of those sentences (compare 2005, 277ff.). However, we must distinguish two versions of this idea. According to the first, the concept of truth in question belongs to the object language. Thus, what is at issue is the equivalence of 'Hunde bellen' and "Hunde bellen' ist wahr', rather than "Hunde bellen' is true'. This is *object language* deflationism, the claim that object language sentences 's' and attributions of *object language* truth to 's' are equivalent. The second idea in play is *metalanguage* deflationism, the claim that object language sentences 's' and attributions of *metalanguage* truth to 's' are equivalent. Here, the equivalence is between 'Hunde bellen' and "Hunde bellen' is true' rather than "Hunde bellen' ist wahr'.

Given this distinction, Patterson argues that while a truth-conditional theory of meaning is compatible with object language deflationism, it is *incompatible* with *metalanguage* deflationism (compare 2005, 280ff.). If, according to metalanguage deflationism, object language sentences 's' and attributions of metalanguage truth to 's' are equivalent, we can fix the meaning of our metalanguage truth-predicate 'is true in German' on the basis of the meanings of the sentences of German. But then, this truth-predicate is of no help in explaining for the meanings of these object language sentences. We cannot *at the same time* use our metalanguage truth-predicate 'is true in German' in such an explanation *and* claim that this predicate can be defined on the basis of the meanings of the object language sentences. Either 'is true in German' explains the meanings of German sentences or the meanings of German sentences explain 'is true in German', but not both.

It might be tempting to assimilate this argument to the Circularity Objection discussed above (section 4.1). But the Metalanguage Objection does not fall prey to the problems raised for the Circularity Objection. It acknowl-

edges that truth has only an expressive role:

> [T]he problem with the claim that deflationism is compatible with including attributions of truth-conditions in an account of meaning as long as these are given only an expressive role is that the deflationist explains this expressive role itself in terms of meaning. (Patterson 2005, 285)

But the Metalanguage Objection fails for a different reason: It relies on a misconception of prosententialism regarding the very idea the objection relies on, the distinction between object language and metalanguage.

Patterson is right that 'is true in German' and the meanings of German sentences cannot simultaneously explain one another. But this problem does not apply to anaphoric deflationism. On this view, metalanguage sentences like "Hunde bellen' is true' *inherit* their meanings from object language utterances like 'Hunde bellen'. But, crucially, *no* object language sentences are involved in *defining* the metalanguage concept of truth. This concept is defined as the truth operator, the linguistic mechanism according to which prosentences can be formed. But the truth operator can be introduced into the metalanguage with appeal to metalanguage sentences alone. Later, it can be applied to every sentence tokening whatsoever, including those one does not initially understand, and including those from other languages. I will come back to this point in the next section. For now, the crucial point is: Prosententialism *does* rely on the meanings of sentences in order to explain the concept of truth, but it only relies on *metalanguage* sentences.

Thus, Patterson's argument boils down to the insight that one cannot specify the truth-conditions of object language sentences unless the metalanguage in which one expresses those truth-conditions is meaningful. This result is a well-known feature of Davidsonian semantics rather than an objection to it. Davidson himself is very aware of this issue and he addresses it when he develops his theory of interpretation into a theory of linguistic

interaction. In any case, this potential problem for Davidsonian semantics is independent from the combination of this theory with prosententialism.

5.3 Cognitive Opacity and the Unity of Truth

Above, I relied on the idea that prosententialism can account for an important fact: We can acquire the concept of truth in our mother tongue and apply it to sentences we do not understand, including sentences from other languages. John Collins calls this phenomenon "cognitive opacity" and argues that it poses serious problems for deflationism (compare 2002, 521ff.). Here, I shall briefly show how prosententialism can overcome these worries. This will also lead to some general remarks about the unity of the concept of truth.

Collins formulates his arguments with a specific version of deflationism in mind, Paul Horwich's 'minimalism' (1990):

> Given Horwich's (1990) deflationism, [...] the meaning of 'true' = the English community's disposition to accept *a priori* the instances of [the scheme "The proposition that p is true iff p"] they can formulate. [...] Upon this basis, Horwich presumes, we can go on to apply 'true' to all other sentences by some form of extrapolation. (Collins 2002, 524)

Collins points out (compare 2002, 524ff.) that the disposition Horwich cites is consistent with practically every conception of truth. And while it may play a role in an account of cognitive opacity, it stands in need of explanation itself. What is needed is some *understanding* of 'true' *in virtue of which* speakers have this disposition. Only this might explain why speakers are able to extrapolate their use of 'true' and apply it to foreign sentences.

> What Horwich requires, then, is an argument that the concept of truth evidenced by the [.] disposition is such as to apply opaquely. Such an argument might be available, but not to Horwich, for it

> presupposes that the possession of the concept of truth is not constituted by the disposition. (Collins 2002, 528)

But such an argument *is* available to another variety of deflationism, prosententialism: If speakers learn how to use 'true', they learn how to form prosentences. That is, they learn how to form sentences which inherit their meanings from other sentence tokenings. Among other things, they thereby acquire a disposition to accept the equivalence of 'p' and 'it is true that p'. However, there is an understanding of the concept of truth underlying this disposition. Namely, this concept is an expressive device, the truth operator, which is defined as the unique anaphoric mechanism for forming prosentences. Whatever anybody else says, if a speaker wants to make an equivalent claim, she can use the concept of truth in order to do so.

This understanding is the basis for many forms of extrapolation. For example, speakers can make 'blind' endorsements of currently unknown sentences like (10) on page 7. Crucially, these function exactly like 'blind' endorsements of currently unknown sentences of *other* languages. This is no explanatory leap since once speakers know how prosentences inherit their meanings – which is tantamount to possessing the concept of truth – all they have to do is uniquely specify the anaphoric antecedent(s) of their prosentences. Thus, anaphoric deflationism can account for cognitive opacity.

This insight also serves a further purpose. Following Tarski, Davidsonian semantics specifies distinct truth-predicates for different languages – 'true in German', 'true in Finnish', ... This has led Davidson to reject the idea that Tarski's or a deflationist approach can account for all there is to truth:

> Tarski has told us much of what we want to know about the concept of truth, and that there must be more. There must be more because there is no indication in Tarski's formal work of what it is that his various truth predicates have in common, and this must be part of the content of the concept (1990, 295).

Collins also stresses this point (compare 2002, 518) and it is a fair one. Whatever it is that the different Tarskian truth predicates have in common, it must be part of the content of the concept of truth. However, the discussion above shows that anaphoric deflationism can neatly account for this problem: The expressive resources demanded from Tarskian 'true in L'-predicates can be met by the truth operator. At the same time, the fact that these predicates are instances of the truth operator explains what they have in common, namely, certain expressive functions. This might not be much, but it is sufficient to account for the unity of the concept of truth.

6   Conclusion

I have argued that Davidsonian semantics and anaphoric deflationism are perfectly compatible and that a number of objections to this claim rely on misconceptions of these theories. This shows that there is hope for combining deflationist conceptions of truth and truth-conditional theories of meaning. However, other varieties of deflationism and other truth-conditional approaches to meaning deserve seperate attention.

Sometimes, the way Davidson himself characterizes the concept of truth already invites a deflationist reading of his position:

> I think of truth as Frank Ramsey thought of probability. [...] Subjective probability is not observable [...]. So Ramsey axiomatized the pattern of preferences of an idealized agent [...]. He stated the conditions on which a pattern of such preferences would be "rational" [...]. Ramsey did not suppose everyone is perfectly rational in the postulated sense, but he did assume that people are nearly enough so [...] for his theory to give a content to the concept of subjective probability [...]. We should think of a theory of truth for a speaker in the same way [...]: both describe structures we can find, with an allowable degree of fitting and fudging, in the

behavior of more or less rational creatures gifted with speech. It is in the fitting and fudging that we give content to the undefined concepts of subjective probability and subjective values – belief and desire, as we briefly call them; and, by way of theories like Tarski's, to the undefined concept of truth (1996, 277f.).

Davidson thinks that the concept of truth can be understood along the lines of its role in a theory intended to make sense of "the behavior of more or less rational creatures gifted with speech", in terms of its role in radical interpretation. But this role has turned out to be intelligible in deflationary terms. Ironically, Davidson's own analogy supports this: Just as Ramsey is concerned with what agents *take* to be probable rather than with what probability *consists* in, Davidson only relies on what interpreters and interpretees take to be *true* rather than on what *truth* consists in. The content truth is given "in the fitting and fudging" is simply its functional, expressive role.

Thus, Davidson has no reason to reject anaphoric deflationism. Moreover, it is natural to frown on his strategy to treat truth as an inexplicable explanatory primitive. John Collins is right:

> [I]t remains unclear just what Davidson's view of truth is, if not a deflationary one. [...] Davidson's principal complaint against deflationism is just that it precludes truth from being the core concept in a theory of meaning. (Collins 2002, 498)

Given that this principal complaint has turned out to be ungrounded and that Davidson already rejects substantial conceptions of truth for independent reasons (1990, 1996), he can happily afford to become an anaphoric deflationist. This would remove some puzzlement about his conception of truth and locate the explanans in Davidsonian semantics where it belongs: in the practice of radical interpretation.[22]

References

Armour-Garb, Bradley, & Beall, J. C. (eds). 2005a. *Deflationary Truth.* Chicago and La Salle: Open Court.

Armour-Garb, Bradley, & Beall, J. C. 2005b. Deflationism: The Basics. *In:* Armour-Garb & Beall (2005a). 1–29.

Bar-on, Dorit, Horisk, Claire, & Lycan, William. 2001. Deflationism, Meaning, and Truth-Conditions. *In:* Armour-Garb & Beall (2005a). 321–343. Originally published in: Philosophical Studies 101, 1–28.

Brandom, Robert. 1984. Reference Explained Away. Anaphoric Reference and Indirect Description. *In:* Armour-Garb & Beall (2005a). 258–281. Originally published in: The Journal of Philosophy 81, 469–492.

Brandom, Robert. 1988. Pragmatism, Phenomenalism, and Truth Talk. *Realism and Antirealism. Midwest Studies in Philosophy*, **12**, 75–93.

Brandom, Robert. 1994. *Making It Explicit. Reason, Representing, and Discursive Commitment.* Cambridge: Harvard University Press.

Brandom, Robert. 1997. Replies to Commentators. *Truth. Philosophical Issues*, **8**, 199–214.

Brandom, Robert. 2002. Expressive versus Explanatory Deflationism about Truth. *In:* Armour-Garb & Beall (2005a). 237–257. Originally published in: Schantz (2002b), 103–119.

Brandom, Robert. 2008. *Between Saying and Doing. Towards an Analytic Pragmatism.* Oxford: Oxford University Press.

Båve, Arvid. 2009. Why is a truth-predicate like a pronoun? *Philosophical Studies*, **145**, 297–310.

Collins, John. 2002. Truth or Meaning? A Question of Priority. *Philosophy and Phenomenological Research*, **65**, 497–536.

Davidson, Donald. 1967. Truth and Meaning. *Synthese*, **17**, 304–323.

Davidson, Donald. 1973. Radical Interpretation. *Dialectica*, **27**, 313–328.

Davidson, Donald. 1974. Belief and the Basis of Meaning. *Synthese*, **27**, 309–323.

Davidson, Donald. 1982. Rational Animals. *Dialectica*, **36**, 317–327.

Davidson, Donald. 1986. A Nice Derangement of Epitaphs. *In:* LePore, Ernest (ed), *Truth and Interpretation. Perspectives on the Philosophy of Donald Davidson.* Oxford: Blackwell. 433–446.

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

Davidson, Donald. 1991. Three Varieties of Knowledge. *In:* Griffiths, A. Phillips (ed), *A. J. Ayer Memorial Essays. Royal Institute of Philosophy Supplement, 30.* Cambridge: Cambridge University Press. 153–166.

Davidson, Donald. 1996. The Folly of Trying to Define Truth. *The Journal of Philosophy*, **93**, 263–278.

Dummett, Michael. 1959. Truth. *Proceedings of the Aristotelian Society*, **59**, 141–162.

Grover, Dorothy. 1977. Inheritors and Paradox. *The Journal of Philosophy*, **74**, 590–604.

Grover, Dorothy, Camp, Joseph, & Belnap, Nuel. 1975. A Prosentential Theory of Truth. *Philosophical Studies*, **27**, 73–125.

Horisk, Claire. 2007. The Expressive Role of Truth in Truth Conditional Semantics. *The Philosophical Quarterly*, **57**, 535–557.

Horisk, Claire. 2008. Truth, Meaning, and Circularity. *Philosophical Studies*, **137**, 269–300.

Horwich, Paul. 1990. *Truth.* Oxford: Blackwell.

Kölbel, Max. 2001. Two Dogmas of Davidsonian Semantics. *The Journal of Philosophy*, **98**, 613–635.

Lance, Mark. 1997. The Significance of Anaphoric Theories of Truth and Reference. *In:* Armour-Garb & Beall (2005a). 282–297. Originally published in: Truth. Philosophical Issues, 8, 180–198.

Löwenstein, David. 2010. Anaphoric Deflationism and Theories of Meaning. *In:* Achourioti, Theodora, Andrade, Edgar, & Staudacher, Marc (eds), *Proceedings of the Amsterdam Graduate Philosophy Conference. Meaning and Truth. Amsterdam, October 1-3, 2009.* Amsterdam: ILLC Publications. 52–66. `http://www.illc.uva.nl/Publications/ResearchReports/X-2010-01.text.pdf`.

Patterson, Douglas. 2005. Deflationism and the Truth Conditional Theory of Meaning. *Philosophical Studies*, **124**, 271–294.

Patterson, Douglas. 2007. On the Determination Argument against Deflationism. *Pacific Philosophical Quarterly*, **88**, 243–250.

Tarski, Alfred. 1944. The Semantic Conception of Truth and the Foundations of Semantics. *Philosophy and Phenomenological Research*, **4**, 341–376.

Williams, Michael. 1999. Meaning and Deflationary Truth. *The Journal of Philosophy*, **96**, 545–564.

Wilson, W. Kent. 1990. Some Reflections On The Prosentential Theory of Truth. *In:* Dunn, J. Michael, & Gupta, Anil (eds), *Truth or consequences: essays in honor of Nuel Belnap.* Dordrecht: Kluwer Academic Publishers. 19–32.