Anaphoric Deflationism and Theories of Meaning

David Löwenstein (Freie Universität Berlin)

`david.loewenstein@fu-berlin.de`

I Introduction

It is widely held that truth and reference play an indispensable explanatory role in theories of meaning. By contrast, so-called deflationists argue that the functions of these concepts are merely expressive and never explanatory. Robert Brandom has proposed both a variety of deflationism—the anaphoric theory—, and a theory of meaning—inferentialism—which doesn't rely on truth or reference. He argues that the anaphoric theory counts against his (chiefly referentialist) rivals in the debate on meaning and thereby paves the way for inferentialism.

In this paper, I give a friendly reconstruction of anaphoric deflationism (section II) and point to a distinguishing feature of the theory with respect to other deflationist proposals. While Brandom simply assumes, but doesn't earn this feature, I propose a natural argument to justify it (section III). Then, however, I point out a subtle but clear sense in which truth and reference can play a role in explaining meaning, even if the anaphoric theory holds. Thus, anaphoric deflationism will turn out to be neutral in the debate on meaning (section IV).

II Anaphoric deflationism

While there is some disagreement over how best to define deflationism, I take it that the most fruitful proposal relies on distinguishing two kinds of questions (compare Armour-Garb and Beall 2005).

THE SUBSTANTIAL QUESTIONS What is truth? What is reference?

THE FUNCTIONAL QUESTIONS What is the conceptual role of truth and the linguistic role of 'true' and related expressions? What is the conceptual role of reference and the linguistic role of 'refers to' and related expressions?

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

Deflationary theories differ in how to make sense of the functional roles of truth and reference. As its name already reveals, the anaphoric theory relies on anaphoric mechanisms. [1] That is, it relies on relations between linguistic expressions that allow some expression tokenings, anaphoric dependents, to anaphorically inherit the meanings of other expression tokenings, anaphoric antecedents. Pronouns are paradigmatic examples for anaphoric expressions. Consider the following example.

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

---

[1]For the original presentation and development of the theory, compare Grover et al. 1975, Brandom 1984, 1988, 1994 and 2002.

Here, 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 anaphoric dependent refers *forth* to that antecedent and thereby inherits its meaning.[2]

II.1   Truth

The main idea of the anaphoric account of truth is to treat complex expressions formed with expressions like 'is true' as *prosentences*. When concerned merely with truth and not with reference, anaphoric deflationism is therefore also often referred to as '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 to, say, sentences or propositions. Rather, truth talk relies on a unique anaphoric prosentence-forming operator.

Consider the following example of ordinary truth talk.

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

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

(3)   Sarah is a maverick.

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.

---

[2]Linguists call the latter phenomenon 'cataphora' and reserve the term 'anaphora' for 'forwards' cases like the former. Cataphora and anaphora are, in turn, subclasses of endophora. Thus, strictly speaking, the theory is endophoric rather than anaphoric.

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

Given the occurrence of the variables both inside and outside of quotation marks, the quantifiers have to be understood substitutionally.

Prosentences are generic. That is, any sentence tokening can be the antecedent of a prosentence, given that it is specified uniquely, e.g., by description or quotation. 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). Prosentences can occur free-standing or embedded in logically complex sentences, e.g., as antecedents of conditionals.[3]

But what are prosentences good for? The answer is that they enable speakers to say things they couldn't say otherwise. For example, one can blindly endorse or reject claims, as in (8), one can reject one from a set of claims without specifying which, like in (9), and one can endorse large or even infinite sets of claims in one breath, as in (10).

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

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

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

Finally, the anaphoric theory also implies an important scheme:

TRUTH SCHEME 'ϕ' is true just in case ϕ.

The left hand side of this biconditional is interpreted as a prosentence, and if the sentence quoted on the left is the one used on the right, they will share their meaning by anaphoric inheritance and thereby validate the biconditional.

II.2   Reference

The main idea of the anaphoric account of reference is to treat complex expressions formed with expressions like 'refers to' as *indirect definite descriptions*. These are anaphorically dependent singular terms which inherit their meanings from other singular term

---

[3]This implies that the Liar is a prima facie admissible prosentence. Grover (1977) has attempted a prosententialist dissolution of the Liar, but I cannot discuss these issues here.

tokenings. Reference talk, the claim goes, does not involve attributions of a relation between linguistic expressions and objects. Rather, reference talk relies on a unique anaphoric pronoun-forming operator.

Take this example to help clarify the idea:

(11) What you referred to as 'that cat' was actually a fox.

Here, the indirect definite description 'what you referred to as 'that cat'' inherits its meaning from the addressee's latest tokening of 'that cat'. Assuming a suitable context, it is therefore a way of saying

(12) What we saw in the garden was actually a fox.

Like prosentences, indirect definite descriptions can be used quantificationally:

(13) All of these names refer to friends of mine.

(14) I referred to Plato earlier.

Like quantified truth talk, (13) and (14) are elaborated as the following substitutionally quantified statements:

(15) For all singular term tokenings 't', if 't' is among these names, then the one referred to by 't' is a friend of mine.

(16) There are some singular term tokenings 't' with the property that (i) I uttered 't' earlier and (ii) the one referred to by 't' is Plato.

What are indirect definite descriptions good for? Just like prosentences, they enrich the expressive resources of a language. Chiefly, they allow for speakers to continue any anaphoric chain of singular term tokenings, as long as the antecedent tokening is specified uniquely. This includes both tokenings of a type-substitution invariant expressions, such as proper names, and expressions not cotypically substitutable, such as pronouns. Furthermore, in the quantificational cases, one may use, say, (13) in order to avoid giving

all the names in question. And one may employ (14) regardless of whether one has used the expression 'Plato' or 'Aristotle's teacher' or anything like that.

Finally, the anaphoric theory also implies an important scheme:

REFERENCE SCHEME 'ϕ' refers to ψ just in case ϕ is ψ.

Changing merely what is focused on rather than its meaning, the left hand side of this scheme is reformulated as an identity statement:

REFERENCE SCHEME* The one referred to by 'ϕ' is ψ just in case ϕ is ψ.

Now, the expression 'the one referred to by 'ϕ'' is interpreted as an indirect definite description which inherits its content from some previous tokening of 'ϕ'. The scheme is thereby validated whenever the 'ϕ' quoted on its left hand side is the one used on its right. Obviously, this includes all type-substitution invariant expressions 'ϕ', e.g., proper names.

III   Anaphoric operators

On the anaphoric theory, prosentences are defined as the unique expression type formed by the truth operator, and indirect definite descriptions as the one formed by the reference operator. How should these anaphoric operators be understood?

III.1   The need for demarcation

Let me elaborate why this question is crucial for proponents of the anaphoric theory.

A standard way to support the idea, against deflationists, that substantive accounts of truth and reference are called for, is to say that *since* 'is true' and 'refers to' are predicates, there have to be underlying properties which substantive accounts of truth and reference need to analyze. Other varieties of deflationism react to this argument by saying that the properties in question are merely deflationary or purely logical. Or they deny that there are properties at all underlying the predicates in question. By contrast, the anaphoric theory is more radical. It denies that truth and reference talk are

predicative at all and holds that it involves unique anaphoric operators instead. This strategy blocks the move from language use to properties from the outset.

Moreover, it also comes in handy elsewhere: Michael Devitt (2002) has argued that deflationism has hidden metaphysical committments. He holds that metaphysics is "explanatory prior" to language (Devitt, 2002, pp. 61, 63), and argues that

> an anti-realist metaphysics is needed to *motivate* the revisionist view of language [...]. If there were not something problematic about the area of reality that 'true' [..] appear[s] to concern why suppose that [it] does not have the standard semantics of a descriptive predicate? (Devitt, 2002, p. 65)

But this complaint is based on the premise that 'is true' is a predicate. Thus, if the anaphoric theory can account for truth talk in terms of an anaphoric operator rather than a predicate, Devitt's criticism misfires. Analogous arguments apply to 'refers to'.

III.2   Defining the operators

I have argued that the anaphoric theory relies heavily on the distinction between the analysis of truth and reference talk as predicative and their analysis as anaphoric. But what accounts for this distinction? Brandom explicitly defines the reference operator and thus distinguishes reference talk from predication. Surprisingly, however, he remains silent about the truth operator. In what follows, I will argue that a suitable definition of the latter can be modeled on the former.[4]

Brandom claims that an expression qualifies as an instance of the reference operator ϱ just in case it satisfies the following criterion solely in virtue of its meaning.

ITERATION CONDITION (REFERENCE) Let $[\phi]$ be the type of ϕ and $([\phi])$ a token of the type $[\phi]$. Then $\varrho([\phi]) = \varrho([\varrho([\phi])])$.

---

[4]In a response to a paper by Mark Lance, Brandom can be read as already hinting at this idea. But even if so, what he explicitly says about this is far from clear (compare Brandom 1997).

Using 'refers' as the paradigm of the reference operator ϱ and stipulating that this alone suffices to pick out some previous utterance uniquely, the criterion states:

(17) the one referred to as 'ϕ' = the one referred to as 'the one referred to as 'ϕ''

According to this definition, expressions like 'the one denoted by 'ϕ'' also qualify as instances of the reference operator. But expressions like 'the one praised as 'ϕ'' or 'the one insulted by saying 'ϕ'' are excluded. For example, one may pick up the meaning of a tokening of the expression 'this important philosopher' by saying 'the one praised as 'this important philosopher''. But saying 'the one praised as 'the one praised as 'this important philosopher''' might well fail to continue this anaphoric chain. Reporting and expressing praise don't coincide.

My proposal for an amendment of prosentialism, the anaphoric account of truth, is strictly analogous. I propose to treat an expression as an instance of the truth operator ϑ just in case it satisfies the following criterion solely in virtue of its meaning.

ITERATION CONDITION (TRUTH) Let $[\phi]$ be the type of ϕ and $([\phi])$ a token of the type $[\phi]$. Then $\vartheta([\phi])$ if and only if $\vartheta([\vartheta([\phi])])$.

Using 'true' as the paradigm of the truth operator ϑ, and stipulating that this alone suffices to pick out some previous utterance uniquely, this second criterion states:

(18) 'ϕ' is true if and only if ''ϕ' is true' is true.

On this definition, expressions such as 'It is the case that ϕ' or, arguably, ''ϕ' is a fact' will turn out to be instances of the truth operator. Other expressions, by contrast, are excluded, e.g., 'It is a pity that ϕ'. For it might well be a pity that my favorite team lost since victory would have been a great reward for a season of hard work. But that doesn't mean that it is a pity that it is a pity that they lost, since their defeat's being a pity doesn't itself have any bad consequences.

III.3   The opposition to predicates

How do these criteria underwrite the claim that truth and reference talk are not predicative?

For a start, the iteration conditions would have to be reinterpreted to fit predicative cases. A truth predicate T, for example, would be defined as holding of any x just in case it also holds of $T(x)$. And a relational reference predicate R would be defined as holding between any x and y just in case it also holds between x and the unique z which stands in R to y. However, even if this can be somehow made to work, the anaphoric theory rules out such an interpretation. Let me show this for the truth operator and keep an analogous argument about the reference operator implicit.

Take three sentences like the following.

(19) Dogs bark.

(20) 'Dogs bark' is true.

(21) "Dogs bark' is true' is true.

On the anaphoric theory, (20) inherits its meaning from (19) and (21) from (20). Thus, they all have the very same meaning. By contrast, if truth talk is understood as predicative, (19) is about dogs, (20) about a sentence about dogs and (21) about a sentence about a sentence about dogs. Thus, there is a difference in meaning after all, a difference one may describe in terms of three levels of semantic ascent, from non-semantic in (19) to semantic in (20) to meta-semantic in (21). On the anaphoric theory, this ascent is an illusion. The semantic 'ladder' is horizontal. Thus, the iteration condition implies that truth talk cannot be predicative.

This last point brings out that the above iteration conditions, together with the rest of the anaphoric theory, don't merely have to be fulfilled solely in virtue of the meanings of the candidate expressions. After all, the expression 'is a quotable expression' arguably also fulfills both of them in virtue of its meaning.[5] Rather, the *resultant* expressions, too, have to share their meanings, as their respective roles in the anaphoric chain ensures.

---

[5]I am indebted to Miguel Hoeltje for confronting me with this example.

This feature is often criticized: Couldn't somebody accept that 'Dogs bark' is true without knowing the meaning of 'Dogs bark'? How, then, are these supposed to share their meanings? But it is entirely possible to accept a prosentence without being able to identify its antecedent. This is even one of the expressive advantages of the truth operator: blind endorsements like (8) on page 55.

IV   Substantivism and theories of meaning

I have argued that proponents of the anaphoric theory can coherently shrug their shoulders at the substantive questions what truth and reference consist in. However, what does this mean for the *answers* actually proposed by substantivists and for explanations of meaning in terms of truth and reference?

IV.1   Substantivism

I take it that substantial accounts in general can be characterized by the claim that there is some substantial property— T, say— had by everything and only what is true, and that whatever is referred to by some expression stands is some substantial relation to it— say, R. Thus, the position can be expressed as follows:

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

'ϕ' refers to ψ just in case 'ϕ' stands in the substantial relation R to ψ.

The first and crucial thing to note about Substantivism is its logical independence from the anaphoric theory. Recall that the anaphoric theory validates the following schemes:

SEMANTIC SCHEMES 'ϕ' is true just in case ϕ.

'ϕ' refers to ψ just in case ϕ is ψ.

On the anaphoric theory, these schemes are true in virtue of the meanings of the expressions formed with 'true' and 'refers'. For example, 'ϕ' is true just in case ϕ because "'ϕ' is true' means that ϕ. Thus, the anaphoric theory implies that Substantivism means Substantivism*:

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

ϕ is ψ just in case 'ϕ' stands in the substantial relation R to ψ.

However, Substantivism* is a claim without any use of truth and reference talk. Thus, it should be clear that it is logically independent from the anaphoric theory.

Mark Lance (1997) has already discovered half of this result, the part concerning truth talk only. He concluded that

> the anaphoric theory is not incompatible with any currently popular account of truth, charitably understood ... [N]o one has yet succeeded even in offering a theory incompatible with the anaphoric account, much less a refutation (Lance, 1997, pp. 283, 297).

However, Lance is too quick with the expression 'charitably understood'. If substantive accounts should be understood as maximally plausible from the point of view of anaphoric deflationism, he is certainly right. But we should also seek a charitable reading of the substantivists' ideas which takes their self-description as actually contradicting deflationism seriously.

What Lance fails to see is that Substantivism turned out to be compatible with anaphoric deflationism because it was stated as a purely extensional claim about expressions like "ϕ' is true' and "ϕ' refers to ψ' rather than about their meanings, or intensions. Full-blown substantivists hold that truth and T on the one hand and reference and R on the other don't merely *coincide*, but that truth *is T* and reference *is R*. In contrast with the above characterization of Substantivism, this idea can be expressed as follows:

FULL-BLOWN SUBSTANTIVISM "ϕ' is true' means that 'ϕ' has the substantial property T.

"ϕ' refers to ψ' means that 'ϕ' stands in the substantial relation R to ψ.

This, however, collides with anaphoric deflationism. The iteration conditions employed to define the truth and reference operators clearly show that "ϕ' has the substantial property T' is not an instance of the truth operator, and that 'the one standing in the substantial relation R to ψ' is not an instance of the reference operator. Even if they

were to fulfill thir respective iteration conditions, the resultant expressions would fail to share their meaning.[6] Thus, these expressions are predicates, the anaphoric operators are not.

Thus, there is a clear disagreement between anaphoric deflationism and substantivism, after all: However the property T and the relation R are understood exactly, substantivists claim and anaphoric deflationists deny that 'true', 'refers' and cognates mean T and R, respectively.

However, this is indeed all there is to the disagreement. In order to reject Full-blown Substantivism, the anaphoric deflationist needs additional arguments against T and R. And in order to reject anaphoric deflationism after Full-blown Substantivism has already fallen, the substantivist needs additional arguments against its functional anaphoric analysis of truth and reference talk. This is where their debate should continue.

IV.2   Meaning

Let me now discuss the implications of these insights for theories of meaning. Brandom thinks that the anaphoric theory makes it impossible to explain meaning in terms of truth or reference. He tries to establish this claim with what I call his Argument from Anaphora:[7]

1. If the anaphoric theory is true, then truth and reference are 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.

---

[6]Colin McGinn (2002) takes up this dialectical position about truth when he claims that the disquotational nature of truth fails to support deflationism, but reveals its nature as "a device of ontological leapfrog". He even *defines* the substantial property of truth by something like the iteration condition.

[7]Compare Brandom 1994 and 2002. The literature on deflationism and meaning is full of cognate arguments. Their shared structure can be made explicit by substituting, say, 'disquotation' for 'anaphora' in the argument presented here.

4. Therefore, if the anaphoric theory is true, then meaning cannot be explained in terms of truth or reference.

However, this argument is far from a refutation of the idea that substantive accounts of truth and reference can be employed to explain meaning. True, the anaphoric theory precludes truth and reference from playing any explanatory role, but it is entirely neutral on what substantivists claim truth and reference *consist in.* If the anaphoric theory is correct, the substantivist has to sacrifice Full-blown Substantivism, but nothing more. When claiming that truth and reference explain meaning, one can simply *give away the words* 'truth' and 'reference', but maintain that *the substantial properties T and R* explain meaning.

Let me spell this out for the referentialist family of substantivist proposals, which Brandom is at most pains to reject. Referentialism, I take it, is the conjunction of a correspondence theory of truth and a causal-historical theory of reference. Thus, referentialists replace 'T' in the above schemes with, say, 'corresponds to a truth-maker' and 'R' with 'is causal-historically connected to'.

Crucially, whenever a referentialist appeals to what she calls 'truth' and 'reference', there will be a clear sense in which she actually appeals to correspondence and causal-historical connectedness.

Traditionally, the situation is this: Truth and reference are supposed to explain meaning, truth is correspondence, and reference is causal-historical connectedness. But explanatory relations are transitive. Thus, at the ground level, it is correspondence and causal-historical connectedness which explain meaning.

By contrast, if the referentialist buys into the anaphoric theory, drops Full-blown Substantivism and contents herself with Substantivism*, this argument becomes superfluous. Rather, the claim that truth and reference explain meaning becomes a potentially misleading shorthand for saying that, *from the outset*, the actual explanans is correspondence and causal-historical connectedness.

Thus, Brandom cannot invoke the anaphoric theory in order to argue against referentialist or other substantivist explanations of meaning. And likewise, one cannot invoke such an explanation of meaning in order to reject the anaphoric theory (like Schantz 2002 and many

others). The anaphoric theory is logically independent from theories of meaning such as referentialism and inferentialism.

V   Conclusion

I have argued, first, that the anaphoric theory can account for its most radically deflationary claim that 'truth' and 'reference' are operators as opposed to predicates. Second, I have shown how substantivists can accept the anaphoric theory if they are willing to give away the words 'truth' and 'reference' and content themselves with the claim that truth and reference coincide with certain substantial properties. Thus, the explanatory value of those substantial properties is untouched by anaphoric deflationism and the debate over the explanation of meaning remains undecided.

However, these results shouldn't be taken to suggest that the debate between deflationism and substantivism is over. Instead, we should continute to discuss the functional analyses of truth and reference talk proposed by deflationists. As for anaphoric deflationism, we are now in a position to assess its distinctive and controversial characteristic, the claim that 'truth' and 'reference' are operators as opposed to predicates. Further, we should discuss the plausibility and explanatory power of Full-blown Substantivism in comparison with its more modest cousin Substantivism*. But, one way or the other, theories of meaning are independent from the outcome of this debate.[8]

References

Armour-Garb, B. and Beall, J. C. (2005). Deflationism: The Basics. In *Deflationary Truth*, pages 1–29.

Brandom, R. (1984). Reference Explained Away. Anaphoric Reference and Indirect Description. In *Deflationary Truth*, pages 258–281. Open Court. Originally published in: The Journal of Philosophy 81, 469–492.

---

[8]For insightful commentaries and helpful criticism, I am grateful to Christian Voigt, Eugen Pissarskoi, Sebastian Bänker, and to the participants of the conference in Amsterdam—especially to David Hunter, who commented on my talk.

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

Brandom, R. (1994). *Making It Explicit. Reason, Representing, and Discursive Commitment.* Harvard University Press, Cambridge.

Brandom, R. (1997). Replies to Commentators. *Truth. Philosophical Issues*, 8:199–214.

Brandom, R. (2002). Expressive versus Explanatory Deflationism about Truth. In *Deflationary Truth*, pages 237–257. Originally published in: Schantz (2002b), 103–119.

Devitt, M. (2002). The Metaphysics of Deflationary Truth. In *What is Truth?*, pages 60–78. Walter de Gruyter.

Grover, D. (1977). Inheritors and Paradox. *The Journal of Philosophy*, 74:590–604.

Grover, D., Camp, J., and Belnap, N. (1975). A Prosentential Theory of Truth. *Philosophical Studies*, 27:73–125.

Lance, M. (1997). The Significance of Anaphoric Theories of Truth and Reference. In *Deflationary Truth*, pages 282–297. Originally published in: Truth. Philosophical Issues, 8, 180–198.

McGinn, C. (2002). The truth about truth. In *What is Truth?*, pages 194–204. Walter de Gruyter.

Schantz, R. (2002). Truth, Meaning, and Reference. In *What is Truth?*, pages 79–99. Walter de Gruyter.