

A DEFLATIONARY THEORY OF REFERENCE

Arvid Båve, Stockholm University

ABSTRACT: The paper first rehearses three deflationary theories of reference, (1) disquotationalism, (2) propositionalism (Horwich), and (3) the anaphoric theory (Brandom), and raises a number of objections against them. It turns out that each corresponds to a closely related theory of truth, and that these are subject to analogous criticisms to a surprisingly high extent. I then present a theory of my own, according to which the schema "That S(t) is about t" and the biconditional "S refers to x iff S says something about x" are exhaustive of the notions of aboutness and reference. An account of the usefulness of "about" is then given, which, I argue, is superior to that of Horwich. I close with a few considerations about how the advertised theory relates to well-known issues of reference, the conclusions of which is (1) that the issues concern reference and aboutness only insofar as the words "about" and "refer" serve to generalise over the claims that are really at issue, (2) that the theory of reference will not settle the issues, and (3) that it follows from (2) that the issues do not concern the nature of aboutness or reference.

Introduction

Many adherents of the deflationary theory of truth have proposed similarly deflationary theories of the notion of reference. The main motivation for these ideas, it seems, is the belief that there is no correct, informative reduction of the form "For all x , x is true iff ... x ..." or "For all x , y , x refers to y iff ... x ... y ...". The literature on deflationism has suffered from considerable terminological confusion, however, so I should make it clear at the outset how I

intend to use the term “deflationism”, and how it can be extended to notions other than the original target, truth.

The definition which accords best with the claims of actual, self-professed deflationists takes deflationism to be the claim that a certain, more or less trivial schema gives an exhaustive theory of truth. The schema may be, e.g., “‘ p ’ is true iff p ” or “(The proposition) that p is true iff p ”. That the schema is *exhaustive* of truth means that it can explain, together with claims not pertaining to truth, everything that a truth-theory should explain (where it may be left open what exactly this is).

The extension to the notion of reference is now fairly obvious: a deflationary theory of reference is one that takes a more or less trivial schema, like “‘ a ’ refers to a ”, to be exhaustive, in the above sense, of reference. Surely, further interesting similarities between these two views could be described, e.g., the disquotational character the theories ascribe to their respective notions. Although I will try to cast doubt on this alleged disquotational character below, a different common feature will be highlighted: that the theories can both give a plausible explanation of the *raison d’être* of their target expressions, which is to increase a language’s expressive power, in a sense that can be made fairly precise.

This paper will not primarily contend with the polemics between deflationism and its rivals, but rather with that between different varieties of deflationism about reference. After having presented and criticised three main types of deflationary theories of reference (sections 1-3), I will in section 4 present, compare, and defend one of my own. The fifth and final section contains a discussion about how the advertised theory relates to broader issues concerning reference.

As it turns out, the different varieties of deflationism about reference that I discern correspond in a strikingly uniform way to parallel deflationary theories of truth, and the main defenders of the relevant theories of reference have also tended to accept analogous theories

of truth. The three broad families of deflationary theories I distinguish are the *disquotational* (Leeds, Field, Horwich), the *propositionalist* (Horwich), and the *anaphoric* (Brandom). As it turns out, the main objections to the deflationary truth-theories can in many cases be transposed rather straightforwardly into arguments against the corresponding theories of reference.

1. The disquotational theory of reference

This theory takes a schema linking an expression, referred to by a quote-name, to its referent, as in “‘*a*’ refers to *a*”, to give an exhaustive account of reference (see Field (1994: 261ff.), Leeds (1978), Horwich (1998a: 118f.)). It thus corresponds to the disquotational theory of truth, according to which the schema “‘*p*’ is true iff *p*” is taken as exhaustive of truth. Horwich proposes a slightly more elaborate version of this theory:

(DR) Tokens of **n** refer, if at all, to *n*,

where the special quotation marks “*” have the effect of disambiguation (1998a: 118f.). While Horwich takes his theory to capture the meaning of “refer”, Field is more interested in redefining these semantic notions. Nevertheless, I will be discussing the disquotational theory only considered as a candidate for analysing our ordinary locutions. The main weakness of the disquotational *truth*-theory from that perspective is that it contradicts the intuition that whether a sentence is true depends in part on what it means. For if the disquotational truth-schema is meant to be exhaustive, then the instances will just come out as true by the meaning of “true”, thus making the truth of, e.g., “Snow is white” depend *merely* on whether snow is white, and not on whether the sentence means that snow is white.

Another problem, which is related to that noted above, is that there is no explanation of

the substitution failure of the right and left hand sides of the schema instances in many contexts, in particular, modal, propositional attitude, subjunctive and counterfactual contexts. To wit, if a disquotational theory of this simple kind were true, the following pairs of sentences should be equivalent, but they clearly are not.

- (1a) It is necessary that $1+1=2$,
- (1b) It is necessary that “ $1+1=2$ ” is true,

- (2a) If we had used “grass” differently, grass might have been white,
- (2b) If we had used “grass” differently, “Grass is white” might have been true,

- (3a) X (justifiably) believes that snow is white
- (3b) X (justifiably) believes that “Snow is white” is true.ⁱ

By devising counter-examples to these equivalences, we see why this objection is related to the first: these sentence pairs fail to be equivalents because what a sentence means is contingent *a posteriori*. Field has proposed various additions to the simple schema in order to accommodate such cases (1994: 272ff.), and further additions pertaining to foreign (by which I mean non-English) sentences and sentences containing context sensitive expressions (1994: §§ 8, 10)). The resulting theory would not be deflationary in our sense, since it is here much more than a single schema that is meant to constitute the theory. More importantly, these proposals, even supposing them to give the right results (see David (1994) for arguments that they do not), could not be conjoined to form an acceptably unified and simple semantics for “true” (see Båve (2006: 93)). Not only do we have separate claims for particular types of sentence contexts, which are blatantly *ad hoc*, it is also unclear how such a theory would deal

with various types of *mixed* contexts, e.g., where modal contexts occur within propositional attitude contexts, where perhaps indexicals and various foreign sentences are also involved, etc. By considering the prospects of a disquotational theory accommodating all these cases, and comparing with the theory which takes propositions, rather than sentences, as primary truth bearers, the choice between them looks obvious (see Båve (2006: 4.3)).ⁱⁱ

I said above that the objections against the deflationary truth theories had analogues pertaining to their parallel theories of reference. In the present case, we can see that if the schema “‘*a*’ refers to *a*” alone is to explain the use and meaning of “refer”, it cannot accommodate the fact that what an expression refers to depends on contingent facts about meaning, baptisms, conventions, etc. Rather, that “Aristotle” refers to Aristotle comes out as true by the meaning of “refer”, which is of course false. This instance of the disquotational reference schema is only *a priori* in the rather stretched sense that *if* one knows that “Aristotle” refers to something, *and* knows the meaning of “refer” and of quote-marks, then one knows that the instance is true. But knowing *that* “Aristotle” refers to Aristotle is a piece of non-vacuous, empirical knowledge of a contingent matter.

There is also an analogue of the objection that the halves of the disquotational truth schema instances cannot be intersubstituted in certain contexts. To avoid irrelevant questions about ambiguity, let us focus at (DR). If this were constitutive and exhaustive of the notion of reference, then “*n*” and “that which tokens of **n** refer to” should be synonymous and thus intersubstitutable in all (but quotational) contexts (see footnote 1). But they are not. For modal contexts, this problem could perhaps be solved by including an “actually” operator in the schema, but what about propositional attitude contexts? To know that the person to which tokens of **a** actually refer is *F* is different from knowing that *a* is *F*. Perhaps further refinements could deal with this problem, but, as with the disquotational truth-theory, it is unlikely that the resulting theory would be acceptably simple and unified.

Horwich also proposes a few further refinements in order to deal with the reference of foreign expressions (that “La Tour Eiffel” refers to the Eiffel Tower), and of indexicals (that “I”, as used by me, refers to me). The refinements go by adding to the theory of reference principles involving the notions of translation and context relativity:

v is the correct translation of $w \rightarrow (x)(v \text{ refers to } x \text{ iff } w \text{ refers to } x)$

$[\text{Int}(w) = *n*] \rightarrow (x)(w \text{ refers to } x \text{ iff } x = n),$

where the function “Int” takes an indexical to a name referring to the same thing (relative to the context in question) (Horwich (1998a: 119f.)).

But note that these principles are not simply drawn from other theories and utilised as auxiliary assumptions in the explanations of the data. They are *included in* the theory of reference. This frustrates the promise to explain every datum of the theory by appeal merely to the schema and claims that are not about reference. For neither of the above principles can be inferred merely from the disquotational reference schema together with claims about translation and indexicals that do not involve the notion of reference. This means that the theory is not deflationary on our definition.

More importantly, the theory is implausible for reasons independent of any deflationist leanings. First, this theory takes the concept of reference to presuppose the concepts of translation and indexicals, whereas it seems clear that one can have the former without the two latter. The claims Horwich includes in the theory, further, are *data to be explained*, so simply including them in the theory seems *ad hoc*. The ideal is to be able to explain these data by a theory of reference which does *not* mention indexicals, in conjunction with a theory of translation and indexicals, respectively. Both Horwich’s propositional theory and my own

accomplish this, and should therefore be preferred.

The comparison between disquotationalism and propositionalism about reference will also make it clear that the culprit responsible for the difficulties in accounting for the reference of indexicals and foreign expressions is the assumption that reference is a relation primarily between *expressions* and objects. This also accounts for its failure to properly account for the dependency of an expression's referring to a certain object upon contingencies of the use and meaning of the expression.

An exactly analogous conclusion can be drawn concerning the failure of disquotational truth-theories to account for the contingency of the truth-conditions of sentences and for the truth-conditions of foreign sentences and sentences containing indexicals. If one takes truth instead to apply primarily to what is said by uttering sentences, and to sentences only derivatively, then these phenomena can be accounted for in a simple and obvious way (see Båve (2006: 4.4)). In a similar vein, a theory of reference according to which it is not strictly speaking expressions that refer to things, but speakers who refer to things by using expressions, avoids the unattractive commitments of the disquotational theory. We will see in greater detail how such a theory should be formulated in section 4.

2. Propositional deflationism about reference

Having refined the disquotational theory as illustrated above, Horwich defines the notion of reference* which applies primarily to propositional constituents:

$$(P) \quad (x)(\langle n \rangle \text{ refers}^* \text{ to } x \text{ iff } n=x),$$

where “ $\langle n \rangle$ ” refers to the propositional constituent expressed by n . He then defines the notion of reference, relating words and objects, in terms of reference*, thus:

(PD) w refers to x iff $(\exists k)(w \text{ expresses } k \ \& \ k \text{ refers}^* \text{ to } x)$.

This theory closely parallels his theory of truth, on which truth is defined as a property of propositions, and *truth* (italicised) as a property that sentences have iff they express propositions that are true (1998b: 134). As in the case of truth, Horwich wishes to leave open the possibility that reference to an object is primarily a property of expressions, and therefore claims that our actual concept of reference is captured *either* by the conjunction of (P) and (PD) (this is the view I call “Propositionalism”), or by the disquotational theory (1998a: 120).

Note that propositionalism avoids the first objections I raised against disquotationalism. With this theory, one can explain how the reference of an expression depends on contingent, *a posteriori* facts about linguistic conventions, baptisms and meaning – simply by taking what propositional constituents are expressed by words to so depend (which is the main project of Horwich’s (1998a) and (2005)).

Secondly, this theory avoids the disquotational theory’s unattractive commitment to include claims about translation or indexicals, and could instead (pending an argument to the contrary) give the right verdicts on these cases by using independent claims from the theories of translation and indexicals. For instance, one can hold that e translates e' iff e and e' express the same propositional constituent. Then, the right consequences concerning reference will simply flow from this independent claim, together with the theory of reference. Likewise, by some similar manoeuvre, one would let the theory of indexicals state which propositional constituents are expressed by various indexical expressions relative to a context, whence their referents relative to a context will be determined by the theory of reference. None of these undertakings are unproblematic, of course. In particular, the problem, noted by Gupta (1993), of deriving generalisations from a totality of schema instances seems to arise here – but at

least the theory gets the explanatory relationships right.

Although propositionalism comes closest to the theory I will propose, it has a number of unattractive features that my own account avoids. Firstly, the commitment to “propositional constituents” is an unnecessary cost. Many, including realists about propositions, would deny their existence or even the meaningfulness of the notion. The cost may at first seem inevitable, for it may be thought that in order to avoid the unacceptable modal and epistemic consequences of reference-disquotationalism (which parallel those of *truth-disquotationalism*), we should propose a theory analogous to *truth-propositionalism*, given how this theory avoids the corresponding problems with truth-disquotationalism. Thus, it may be thought, we must speak of entities that stand to referring expressions the way sentences stand to propositions. Since referring expressions are constituents of sentences, these would be the propositional constituents. But we do not need to avoid the problems of reference-disquotationalism thus “analogously”, for there is another option: the theory presented in section 4, which steers between the commitment to “propositional constituents” and the consequences of disquotationalism.

The trick is to take reference to be a relation between speakers and things (exactly how this idea can be incorporated in a recognisably *deflationist* theory will be seen below). Besides avoiding the problems with disquotationalism, this idea accords with actual use. Since Horwich aspires to give a theory about “refer” as we actually use it, the fact that he does not take it to relate speakers to things is a second drawback.

Secondly, the proposition that Plato was Greek, say, and other propositions of the same form, are commonly taken as *singular* propositions, i.e., propositions such that the constituent corresponding to the name is the object referred to (note that this is weaker than the view that co-referring names can be intersubstituted in intensional contexts). But then, Horwich’s view would entail that in these cases (which are the paradigmatic cases of

reference), reference is a relation between an object and itself! Horwich is careful to qualify that the propositional constituents are “*de dicto* propositional constituents” (1998: 120). But the very idea that there are *de dicto* propositional constituents corresponding to names contradicts the widespread Kripkean view that sentences like “Plato was Greek” only express singular propositions, and that names have no meanings beyond their referents. Whether Horwich makes that qualification or not, then, his view comes with considerable costs.

Thirdly, and most importantly, the notion of a propositional constituent and the angled brackets used to refer to them are technical and unfamiliar to normal speakers, which, I will argue, poses serious problems for his theory. It is by no means wrong *per se* to use technical notions in a meta-language that states a semantic theory for a piece of natural language. Recall that I am discussing the theory of reference taken as a semantic theory for a small portion of a natural language. By this, I mean *not* that the theory of reference can be a stipulative definition of “refer” (possibly diverging from the ordinary meaning of the word), which may then be used in a semantic theory to characterise the meaning of the expressions of a language. This is rather the way things work in certain truth-theoretic semantic theories. If the theory had that purpose, the technicality would not be unacceptable, since the only relevant requirement would be that of explanatory usefulness, not accordance with actual use. What I mean, rather, is that the deflationary theory of reference is a characterisation of the word “refer” as used by English speakers, where that characterisation will be made using meaning-theoretic concepts available to deflationists about truth and reference (and I am assuming with the majority of philosophers that these will not include truth and reference, respectively). It is also in this sense that I am assessing Horwich’s aspiration to give the actual meaning of “refer”, and it is from that perspective that the technical notation appearing in (P) becomes a problem.

First, a minor point *ad hominem*. Horwich’s meaning theory is a use-theory on which

the notion of speakers' disposition to accept sentences plays a central role. Although he never states this explicitly in the passage about "refer", he is committed to saying that speakers' meaning what they do by "refer" consists in the fact that their underived disposition to accept (P) is the explanatorily basic fact of their use of the word. (This also follows on the plausible assumption that "refer" should stand to (P) as "true" stands to (PS) on Horwich's account of the meaning of "true".) But, of course, speakers are not disposed to accept (P), since they have no idea what the brackets mean. Nor would they have any disposition to accept some paraphrase of (P) using "propositional constituent".

The real problem with (P) as a characterisation of the meaning of "refer", however, is independent of any specific theory of meaning. No matter *how* (P) is supposed to characterise the meaning of "refer" – perhaps it is taken to describe some inner, neural goings-on, or define the *concept* of reference (or whatnot) – it cannot be denied that it defines reference as a relation between propositional constituents and things. So if our word "refer" stands for that relation, or expresses that concept, or is the outer counterpart of that mentalese predicate (or whatnot), then the ordinary uses of "refer" should relate propositional constituents to things. But then, at least some of the sentences of the form "x refers to y" that English speakers use should be such that the left slot is occupied by some expression referring to propositional constituents.ⁱⁱⁱ But, of course, none are, for there are no such expressions in natural languages. This argument goes through even if we grant the existence of propositional constituents. The problem is not the mere fact that the notion of a propositional constituent is used, but the fact that it is used in such a way as to pose unreasonable demands on speakers.

Horwich could bite the bullet here, and claim that since there are no ordinary words for propositional constituents, this just means that all ordinary uses of "refer" are *derivative*. First, it is of course implausible to say, for any expression, that all of its uses are derivative; a sound methodological rule of thumb is to assume that the most common use is the primary one. The

idea that they are using the word derivatively means that when they utter a sentence “*a* refers to *b*”, this is to be understood as “really meaning” that *a* is related to some propositional constituent that refers to *b* (or some other proposition similarly involving the concept of a propositional constituent). We may disregard what exactly the “really meaning” comes to. The point is that speakers, on this hypothesis, are using “refer” *loosely* or to *abbreviate* (that is, to abbreviate that which is really meant).

But the idea that speakers *really mean* that *a* is related to some propositional constituent that refers to *b* (or some such) just sounds implausible. People do not “really mean” such technical things. On the hypothesis under scrutiny, speakers would *implicitly* be taking the relation to hold between propositional constituents and things, but explicitly speak as if it relates other things. But that they are *implicitly* taking the relation to hold between propositional constituents and things entails that they are operating with the concept of a propositional constituent.

Perhaps one could swallow this, arguing that the operations are *unconscious*. However, what can be gathered from paradigmatic cases of loose or derivative talk is rather that the concepts involved in that which is really meant must be consciously grasped by the speaker. If we ask a speaker using “refer” something like, “To be precise, do you mean that ... propositional constituent...?”, we would of course get a negative answer. This type of evidence is commonplace in pragmatics, and there is no reason to say that this case should be treated differently.

The idea about implicit, unconscious grasping of technical concepts is starting to look rather gratuitous, like the idea that there is an invisible fairy circling the Earth – difficult to disprove conclusively, but not something we want to commit ourselves to. So, everything else being equal, we should opt for the theory to be defended below, which comes without this cost. In fact, I think I have already taken the suggestion too seriously. For what is the relevant

difference between (P) and the absurd suggestion that it is not primarily people who know things, but abstract objects called “*sapientia*”, which are very different from people? We *speak* as if it is people who know things, but if my imaginary interlocutor is right, it may well be that underlying this use is the idea that it is *sapientia* that know. Admittedly, we say that things other than people “refer” more often than with “know”, but, as we will see in section 4, these uses are most plausibly taken as derivative.

3. The anaphoric theory of reference

The truth-theoretic cousin of this theory of reference is the Prosentential theory of truth, first presented by Dorothy Grover *et al.* (1975), and developed by Robert Brandom (1994, 2002). Brandom is also the originator of the anaphoric theory of reference. Neither of these theories qualifies as deflationary in any obvious way, since neither operates with a schema. However, they have both been so categorised (also by Brandom) and I will follow this practice here without further ado.

On Brandom’s view, “is true” is not a predicate, but a “prosentence-forming operator”, i.e., an expression which can join with other expressions to form a prosentence, e.g., “That is true”. The semantics of these expressions is to be understood in analogy with pronouns (and other anaphoric expressions). In sentences containing what Geach called “lazy pronouns”, like “If Mary is home, then *she* is sick”, where “she” has “Mary” as its antecedent, the pronoun will simply have whatever content the antecedent has. A lazy prosentence, like “That is true” now stands to *its* antecedent, which, by analogy, will be a *sentence*, in the same way that “she” stands to “Mary”. Thus, if uttered in response to “Snow is white”, the prosentence simply inherits its content, whence the utterance will effect an assertion that snow is white.

Not all anaphoric expressions can be explained on these lines, however. For instance, in “Every number is such that multiplying *it* with 2 yields an even number”, “it” clearly cannot

be replaced by its antecedent, “Every number”, why they cannot have the same content. Rather, “it” should here be conceived of as a bound variable. Similarly, say the prosententialists, the sentence “Everything he said is true” should be understood as having the logical form of “Everything he said is such that if he said that it is true, then it is true”, where “it is true” in both occurrences are semantically atomic prosentences, functioning, by analogy with pronouns, as bound *propositional* variables. This analysis is supposed to vindicate the Ramseyan paraphrase “For all p , if he said that p , then p ” (Grover *et al.* (75ff.)). The prosententialists then devise various further ideas for handling other occurrences of “true”, but I will here rest content with this brief description (see Båve (forthcoming) for a more detailed, critical discussion).

Brandom’s anaphoric theory of reference begins with an analysis of expressions of the form

(t) the F (the speaker) referred to as t ,

where “ t ” stands proxy for a *name of* a linguistic expression, like ““Leibniz”” (note the double quote-marks), and “ F ” stands for a general term, like “thing”, “man”, etc. An example is “the mechanic Joe referred to as ‘that airhead’” (1994: 311), or simply “the person referred to as ‘Leibniz’” (1994: 318). These expressions are called “anaphorically indirect descriptions”, and their semantic functioning, says Brandom, is *not* that of the superficially similar definite descriptions. Rather, they are anaphoric expressions, like “it”, or “the senator” in the sentence “A senator stood up, ..., and the senator spoke”, where “the senator” behaves like “he”, having “A senator” as antecedent (1994: 309).

Likewise, expressions of the form (t) are anaphoric dependents, but are special in that their antecedents are specified by a phrase *inside* them. Thus, the indirect description “the

mechanic Joe referred to as ‘that airhead’” is an anaphoric expression whose antecedent is Joe’s utterance of the expression “that airhead”. Thanks to that specification, the whole expression thus works just as “he”, as uttered immediately after Joe’s utterance containing “that airhead” (1994: 312). The point of having such an expression, Brandom says, is that when one fails to recall someone’s name, and there is no expression recently uttered to which a simple “he” might anaphorically anchor, one may nevertheless refer to him by specifying, with an indirect description, a long past utterance that may serve as antecedent.

Inventive as this is, it would of course be more natural to treat these expressions just like superficially similar expressions, like “the one Joe startled by his remark about airheads” (1994: 313). Unfortunately, Brandom’s reply to this objection is not so much a reply as it is a set of criteria for circumscribing the class of expressions he wants to single out for special treatment (i.e., the anaphorically indirect descriptions). The most important criterion is that an *iteration* of the expression be intersubstitutable with the original, where an iteration of “What *t* refers to” is “What ‘what *t* refers to’ refers to”. But even if no expression except what Brandom calls the anaphorically indirect descriptions satisfies this criterion, this does not seem to be enough reason to claim that the latter form a distinct semantic category. In fact, any theory of reference which takes it to be an ordinary binary predicate and which validates the unproblematic instances of “‘*a*’ refers to *a*” will guarantee this intersubstitutability and thus provide a refutation of such an argument. Also, the account Brandom gives of the usefulness of “refer” seems equally open to a theory with the more obvious grammatical analysis.

Most of the arguments that have been levelled against Prosententialism concern its grammatical analyses of various sentences (see Båve (forthcoming), Wilson (1990), and Kirkham (1992: 325ff.)). This seems to be the main weakness of Brandom’s theory of reference as well. Besides the counterintuitive analysis of expressions of the form (t), which

forms the basis of the anaphoric theory of reference, Brandom proposes that every other sentence containing “refer” and other referential locutions be “wrestled” (as Brandom says on p. 307) into sentences containing expressions of the form (t), that is, that their semantic functioning is explained by paraphrases containing expressions of the form (t). Just as the denial that (t) expressions work like ordinary descriptions, this idea naturally posits a rather complex and surprising semantic structure in sentences that appear to belong to more familiar categories.

Note that if the claim were merely that the paraphrases are equivalent, then the theory would be incomplete, since this would fall short of giving an *explanation* of the semantic functioning of the target sentences. For the theory to be explanatory, the paraphrases must be taken to involve a stronger claim, e.g., that the paraphrases display the true logical forms of the target sentences. This is what makes this theory contentious.

While the above weakness concerns grammatical analysis, others pertain more directly to semantics, and consist in the theory’s yielding counter-intuitive results about which sentences are equivalent, the modal status of sentences, etc. To begin with, the sentence “The term ‘Leibniz’ denotes Leibniz” is paraphrased as “The one denoted by the term ‘Leibniz’ = Leibniz” (1994: 316). In the latter, the whole phrase to the left of “=” has “Leibniz” as antecedent. But now consider “If Leibniz is *F*, then he is *F*”. This sentence expresses a necessary, *a priori* proposition. Thus, the content of “he” simply is that of “Leibniz”, just as it should be. But if the indirect descriptions are to work the same way, then the above sentence ought to have the same content as “Leibniz = Leibniz”. But neither “The term ‘Leibniz’ denotes Leibniz”, nor “The one denoted by the term ‘Leibniz’ = Leibniz” is necessary or *a priori*. The anaphoric theory thus shares the disquotational theory’s implausible consequence of making referential relations between expressions and objects necessary *a priori*. (As noted in section 1, these sentences are only *a priori* in a “stretched” sense.) Exactly the same

problem afflicts Prosententialism, which takes “‘Snow is white’ is true” to have the same content as “Snow is white” in virtue of the truth-locution’s having the latter sentence as its anaphoric antecedent (see Båve (forthcoming)).

Brandom does not mention the epistemic status of reference-claims, but he does try to account for their contingency. The strategy is to show that some sentences of the form “The one referred to as ‘Leibniz’ in w is not Leibniz” are true. This is so, Brandom says, because “the candidate antecedents of [the indirect description] are tokenings of the type ‘Leibniz’ that are uttered in w ” (1994: 318). But how can an anaphoric expression, used in the actual world, have a *non-actual* tokening as its antecedent?

Even granting that this cross-world cross-reference is intelligible, however, it seems clear that no paradigmatic anaphor in any natural language display this behaviour. We can say, “In non-actual world w , Plato did not become a philosopher, but he was still a man”, but the antecedent of “he” is of course actual here. If paradigmatic anaphors never cross-refer to non-actual expressions, then “refer” simply does not function like paradigmatic anaphors, which undermines the very idea that “refer” somehow *is* an anaphoric expression.

There is an analogue to this objection in the case of Prosententialism: in order to deal with “Goldbach’s conjecture is true”, Brandom says that its antecedent is the sentence that “Goldbach’s conjecture” *denotes*, i.e., “Any even number is the sum of two primes” (1994: 304). But there is no paradigmatic anaphor in natural language such that it is the *denotation* of its antecedent that serves as the anaphor’s substituend. So, if no other solution is possible, then the idea that “is true” is somehow anaphoric is undermined.

In conclusion, the contingency of reference relations has not been satisfactorily dealt with. What is worse, however, is that even if the account of world-relativising sentences would work, it would seem to *contradict* the theory’s consequence that “The term ‘Leibniz’ denotes Leibniz” has the same content as “Leibniz = Leibniz”. For the former result cannot be

used to show *why we were wrong* to conclude that these sentences are strictly equivalent.

The most serious criticism of these anaphoric theories, I believe, is that the comparison with paradigmatic anaphors is simply idle. There are a number of linguistic data that Brandom's theory of reference accommodates, e.g., the fact that expressions of the form "*a*" and the corresponding "the one referred to by '*a*'" are intersubstitutable in extensional contexts. But why should this be expressed by using the locutions "anaphor", "antecedent", etc., instead of simply saying that they are intersubstitutable? What is the point of expressing these facts in anaphoric terms?

We can get clearer on this by noting, first, that anaphoric phenomena must have some underlying semantic explanation, perhaps in terms of denotation, perhaps, as Brandom would have it, in terms of substitution-inferences (1994: Ch. 7). But regardless of how these notions of "anaphor", "antecedent", etc., are explained, the anaphoric theory about "refer" will then have an equivalent theory that only uses generic semantic notions like "equivalent", "intersubstitutable", etc. Given a complete account of anaphors, one could deduce this theory from the anaphoric account (if the latter is clear and coherent enough).

In view of Brandom's own account of anaphors, his theory of reference does little more than claim that expressions of the form "*a*" and the corresponding "the one referred to by '*a*'" are intersubstitutable. It is thus essentially this claim, together with the paraphrases of other sentences containing "refer" into sentences containing expressions of the form (t), that constitutes the theory. Since it is clear that this theory need not mention anaphors at all, we must conclude that the appeal to anaphora is idle. Therefore, "refer" has no interesting connections with anaphoric expressions.

Further, if indeed this theory comes to no more than the claim that "*a*" and the corresponding "the one referred to by '*a*'" are intersubstitutable, plus the paraphrases, it is hard to see why we should not instead go with the schema "'*a*' refers to *a*", and avoid the

contentious claims about the logical form of sentences containing “refer”.

The same critique can be levelled at Prosententialism: instead of accounting for the equivalence between “That snow is white is true” and “Snow is white” by saying that the former is a prosentence with the latter as antecedent, one might as well simply say that they are equivalent (see Båve (forthcoming) for a detailed criticism on these lines), and similarly for other prosentential analyses of sentences containing “true”.

The only case, I believe, that could be made for saying that “true” and “refer” are somehow essentially anaphoric in nature would be a demonstration that they share a number of idiosyncratic or surprising features with paradigmatic anaphors. This would suggest an interesting connection between these classes of expressions, perhaps that competence with them is realised by the same specific part of our language faculty. But given the awkward grammatical analyses, the counter-intuitive consequences concerning equivalence, and the positing of cross-world anaphoric relations, what emerges is that they do not share surprising, idiosyncratic features, since they do not share enough features at all. The fact that both can be characterised in broad terms such as “equivalence” and “consequence” does not amount to an interesting connection, since this is the case for many other expressions.

4. A theory of our actual concept of reference

Paul Horwich notes that

in ordinary language the term most often employed in connection with [Horwich’s deflationary] reference concepts is “about”, rather than “refers”. We tend to speak of *people* as referring to things, and of them doing so in virtue of their words being about those things. (1998a: 121 – original emphasis)

One may wonder, then, why he proposes a rather recalcitrant theory as a theory of *the meaning* of “refer”. In any case, I will now present a theory which takes ordinary language

seriously. Hereby, the theory will not only accord better with intuition, but also yield a better account of the usefulness of these locutions, and, perhaps most importantly, yield a new conception of the broader philosophical differences between deflationists and its rivals.

Recall that a deflationary theory is one which takes a more or less trivial schema (or a claim about it) to be exhaustive of the notion in question. On the account I will give, we should first analyse the expression “about”, and then explain “refer” in terms of it. The schema I hold to be implicitly defining “about” is:

(A) That $S(t)$ is about t ,

where $S()$ is a sentence context with a slot for singular terms. There might be some disagreement about what expressions may substitute “ t ”, but I am inclined to include all singular terms, as well as natural kind terms, abstract nouns, and plurals.^{iv} It does seem right to say that (the belief) that water is wet is about water and that (the belief) that love hurts is about love, and that (the belief) that horses are fast is about horses. Note that on this theory, aboutness, the relation expressed by the binary predicate “is about”, is a relation between propositions and objects, rather than, as Horwich intimates in the quote above, between words and objects.

The intent of (A) is that all its instances are to be regarded as analytic or unconditionally assertible in virtue of the meaning of “about”, or, if you will, as implicitly defining “about”. This can be spelled out in many ways. For instance, on Horwich’s use-theory, the claim would be that the meaning of “about” is constituted by the fact that speakers have an underived disposition to accept the instances of (A) (and that this constitutes its meaning because it is the explanatorily basic fact about its use). But I will not venture into these issues here. Note, however, that (A) is not only not committed to propositional constituents, it is not

by itself committed to propositions either, but only to “that”-clauses, which trivially exist.

The instances of (A), like “That Quine is a philosopher is about Quine”, may at first appear a little awkward. But by looking at certain other sentences containing “about”, we can alleviate the impression that they are ungrammatical or semantically anomalous. Take the well-formed “He believes something about John”. Now, one might respond to an utterance hereof by asking, “What does he believe?”, and the reply might be “That John is tall”. This “something”, which is “about John”, then, must be “that John is tall”. So it ought not be incorrect to say, “That John is tall is about John”.

Of course, we would seldom, or never, utter an instance of (A), but this does not show that it does not implicitly define “about”. In fact, if it does, there is a pragmatic reason why we shouldn’t utter it – it would be trivial and useless. Ordinary speakers also find instances of the propositional truth-schema strange, but can be taught to see that it is well-formed by pedagogical explanations like the one above.

Note that, in order to square with (A), “He believes something about John” must be seen as a contraction of “He believes something *which is* about John”. Further additional assumptions may be needed to deal with other occurrences of “about”. For instance, “I dreamt about x ” might be taken as abbreviating “I dreamt *something which was* about x ”. We would probably also need to add the schema “Whether $S(t)$ is about t ” in order to handle “He was asking (something) about you”. In order to deal with the expression “I speak about x ”, further, we could assume that “speak” simply means *say something*, or *say things*, in which case this would be but another example of the same kind as “He believes something (or things) about John”.

I should be the first to say that (A) does not by itself explain all uses of “about”. That is, for some cases, some additional pragmatic treatment may be needed. Perhaps we even need to reckon with additional senses of “about”. In that case, however, I submit that the one captured

by (A) is the philosophically interesting one. Here is a case in point: if two speakers are uttering sentences like “The Conservatives will win”, “No, the Radicals will win”, etc., it would be reasonable to say that they are *talking about politics*, even if neither has said that *politics* is such and such. But this seems to be a species of “loose talk”. Suppose *A* says that the Conservatives will win. *B* now says, “*A* said something about politics”. Although it would be pedantic to the extreme, it would still be reasonable from a semantic point of view to object, “Well, *A* didn’t strictly speaking say anything about politics, he said something about the Conservatives”. To say something about politics *strictly speaking* is to say, e.g., that politics is a dirty business. Also, for some reason, it seems even more plausible to say that one *makes a statement about politics* only if one says that politics is so and so, and not if one merely says, e.g., that the Conservatives will win. To take another example, suppose someone is asked to “say something about politics”. Surely, saying, “Politics is a dirty business”, rather than, “The Conservatives will win”, is to take the demand more literally. Thus, (A) seems to cover the literal sense of “about”, which underlies the various loose uses speakers may make of it.

Here is a test to see whether a proposition is about a thing in the strict sense: if a proposition is about *x* in the strict sense, then it is true to say that concerning *x*, it is the proposition that ... he/she/it/they ... We can see that the proposition, call it “P”, that politics is a dirty business passes the test, because concerning politics, P is the proposition that *it* is a dirty business. Nothing such can be said about the proposition that the Conservatives will win. It *might* also be that “about” has an additional sense on which the latter proposition is strictly speaking about politics (although I doubt it), but this test seems to capture the philosophically central sense of the word.

It is important to keep in mind that it is primarily propositions, not sentences, that are about things. Negligence of this fact may make it seem that existential quantifiers should be

among the substituends of “t” in (A). It may seem, for instance, that “One of the candidates wears brightly coloured suits” is about Hillary Clinton, if she happens to be the only candidate wearing brightly coloured suits. But if this sentence is about Clinton, then this is so only in a derivative sense, i.e., by saying something about her. But it may be that, at least in certain contexts, this sentence *does* say something about her, namely that Clinton wears brightly coloured suits. Similarly, I might say something about Jones by saying “A certain philosopher (wink, wink) is here”, namely, that Jones is here. But this, I believe, is not quite relevant for the theory of aboutness, since it has to do rather with the fact that one can sometimes communicate a singular proposition by uttering an existential one. *A fortiori*, this is no reason to believe that (the belief) that something is *F* is about some thing, if that thing is the only *F*. Also, the fact that a sentence saying that a sentence is about something may be true or false depending on the context should not be taken to indicate that aboutness is contextual. Rather, for a proposition that is (context-independently) about an object, it may be contextual whether a given sentence (as used in a given context) expresses that proposition or not.

The notion of reference is now given by:

(R) *a* refers to *b* iff *a* says something (which is) about *b*,

where “*a*” always stands for a *speaker*. One might now, firstly, ask whether this is really a *deflationary* theory of reference, given how it defines it in terms of another notion. To me, it looks recognisably deflationary in spirit, especially compared to theories that propose naturalistic reductions of reference, etc. One might also argue that given the tight connection between “refer” and “about”, and given the clearly deflationary character of the latter, the former notion inherits the deflated nature of the former. The predicament is shared by Horwich, on whose treatment of sentence-truth, a sentence is true iff it expresses a proposition

that is true (where “true” in the latter occurrence is given a deflationary analysis). The same holds for his account of expression reference, on which an expression refers to x iff it expresses a propositional constituent that refers (in the deflationary sense) to x . Just as with prosententialism and the anaphoric theory of reference, it is not obvious how to define “deflationism” so as to include this theory of reference.

Here is a stronger case for calling (R) deflationary, given the deflationary account of “about”. Consider:

(R') a refers to b iff there is an $S()$ such that a says that $S(b)$,

where we quantify into sentence context position (perhaps this could just be seen as ordinary predicate quantification, but I will not enter into such technicalities here). Now, on the assumption of (A), it seems that (R) and (R') say more or less the same thing. And since (R') is more clearly deflationary according to our definition, so is (R).

A major advantage of (R) is its fit with the normal use of “refer”, in that it takes reference to be a relation between a speaker and an object. The philosophical usage, which takes reference to hold between expressions and objects, I believe, should be seen as a something I will call *attributive ellipsis*. To wit, “ e refers to a ” is an attributive ellipsis for (something like) “ e is used (by the speakers in question) to refer to a ”. By (R), for e to be used to refer to a is for it to be used to say something about a . Of course, one normally cannot say something *merely* by uttering a name, wherefore, of course, the phrase “ e is used by speakers to say ...” should not be interpreted thus uncharitably. Rather, by “ e is used to say”, we mean something like, “is used *in sentences* so as to say”. (Similarly, one can use a brush to make portraits, which does not entail that you do not also need paint and a canvas. This example just shows that the above formulation, “is used to ...” is admissible despite the fact that one cannot say

something merely by uttering a name.)

In my opinion, the phenomenon I am calling attributive ellipsis is pragmatic (and should thus be distinguished from the various syntactic phenomena called “ellipsis”, like “sluicing” and “VP-ellipsis”, etc.). I will proceed by first calling attention to a rather different example, offer a treatment thereof (that should be fairly uncontroversial), and then go on to claim that “*e* refers to *a*” should be explained the same way.

In discussions about perception and *qualia*, philosophers (including myself) have spoken of visual experiences as “being green/red/etc.”. A sentence ascribing a colour to an experience, however, is unintelligible if taken as literally, or, in other words, semantically anomalous (but of course grammatical). Still, they are used, seriously and sincerely. The explanation of this linguistic behaviour therefore has to be pragmatic, i.e., an explanation to the effect that what is meant, and what the audience takes the speaker to mean, is distinct from what is semantically encoded in the sentence, and pragmatics simply is the science of this type of communication/interpretation. By uncontroversial pragmatic principles, what is meant should (1) be reasonable (unlike the category mistake) and (2) be something that lies close to the literal meaning of the sentence uttered. In the case of “Experience *e* is green”, this is plausibly that which is expressed by “*e* is an experience *of* something green” (or some such).

Attributive ellipsis, then, occurs when someone utters a sentence which involves an anomalous attribution to something, such that the pragmatic interpretation explaining (away) the anomaly takes that which has been communicated to be something that could be semantically expressed only by a “longer” sentence. By “longer” is actually meant one which is like the one uttered, but contains a few words in addition – hence “ellipsis”.

If I am right that “*e* refers to *x*” is semantically anomalous, then a plausible candidate for what we mean by such sentences, following the model above, is (something like) the

proposition that e is used (by the relevant class of speakers) to refer to x .^v Since this is a pragmatic phenomenon, some indeterminacy of what is communicated is to be expected – hence “something like”. A further concession: there should be no qualms about my *assuming* that “ e refers to x ” is anomalous, since I am not here giving an argument for this claim, but rather trying to provide an explanation *on the basis of my theory* of a superficially recalcitrant fact of usage. Given (R), this is simply a plausible candidate, in that it lies close to the structure and content of the sentence uttered, and, first and foremost, it makes good sense. By (R), “ e is used (by speakers) to refer to a ” in turn means the same as “ e is used (by speakers) to say things about a ” (the “things” being, e.g., that a is tall, that a is short, and so on). On this model, we can give a plausible account of how the technical use of “refer” grew naturally out of the ordinary usage and of how it is that philosophy students can grasp this usage so quickly without being presented with any explicit definition. We can hereby also give an account of “refer” as unambiguous in both types of usage, thereby satisfying Grice’s razor: Do not multiply senses beyond necessity!

Our model also fits neatly with intuitions about the reference of indexicals. For instance, take the intuition that “I”, if used by a , refers to a . Now, that a *uses* “I” means that a utters a sentence containing “I”. Let the sentence be “ $S(I)$ ”. By the theory of indexicals, a will thereby say that $S(a)$, which, by (A) is about a . The reference of foreign expressions can be treated in a similar fashion. Take the intuition that “La Tour Eiffel” refers (in French) to the Eiffel Tower. An independently plausible claim is that “La Tour Eiffel” can be used, i.e., that sentences containing it can be uttered, to say (in French) *that* the Eiffel Tower is large, that the Eiffel Tower is small, and so on, and all of these things are *about* the Eiffel Tower, by (A). Like Horwich’s propositionalism, we do not need to include special clauses concerning indexicals or foreign expressions in the theory of “about” and “refer”.

There are of course many derivative uses of “refer” that are not exclusively philosophical

and do not attribute reference to expressions. Someone may say, for instance “This refers to Moses”, indicating vaguely some passage in a text. Note that one might just as naturally say of the same “thing”, the passage, that it *says that* so and so. This example also seems to fit with (R), for it seems that instead of “This refers to Moses”, one might with no loss have said, “This says something about Moses”. The reasonable interpretation is something like, “By (way of) *this* (i.e., the passage), someone (or, the person in question) says that *p*/refers to *x*”.

Another example may be “The war was about *x*”. One proposal is to take this to be spelled out as, “The proposition such that disagreement whether it is true lead to the war is about *x*”. This seems to be more or less what we mean by “The war was about *x*”. The proposition in question could be, e.g., that slavery should be abolished, which, by (A), is about slavery.

There is an intimately related deflationary theory of truth, on which the schema “That *p* is true iff *p*” is taken as exhaustive of truth. This can then be augmented by the claim that sentences of the form “*s* is true”, where truth is (superficially) ascribed to sentences, are treated as elliptic for something like “What *s* says is true” or “What *s* is used to say is true”. This theory similarly explains the dependence of sentence-truth upon context and language by reference to independently plausible claims (that “Schnee ist weiss” says (in German) that snow is white, and that “I am hungry”, uttered by *a*, says that *a* is hungry, and so on (see Båve (2006: 4.4) for a detailed account). Compare the elliptic phrase “believe him”: since persons cannot strictly speaking be believed, we are guided by pragmatic principles to interpret this as an ellipsis, to be spelled out as something like “believe what he said”.

Finally, what is the point of having a word implicitly defined by the schema (A)? Here, again, there is a tight connection with the notion of truth, in that the common purpose is, in a specific sense, to increase the expressive power of a language. To get clearer on this notion, let us first define $I(s, L)$ as the set of all and only the sentences of L which are entailments of

s. Now, we can define:

$$(EP) \quad \text{An expression } e \text{ increases the expressive power of a language } L \equiv_{df} \\ (\exists s_1 \in L \cup \{e\})(\forall s_2 \in L) (I(s_1, L) \neq I(s_2, L))$$

More colloquially, we could say that e increases the expressive power of L iff there is a sentence containing e and expressions of L which has a set of entailments that do not contain e , which is not the entailment set of any sentence of L . We can now see that “true” increases the expressive power of a language (not containing propositional quantifiers or synonyms to “true”) by considering the sentence “Everything he says is true”. This sentence entails every sentence of the form “If he said that p , then p ”, which no sentence of the original language (given the above restrictions) does (cf. Båve (2006: 29ff.)).

Now consider “I believe everything he said about Quine”. Whereas the sentence “I believe everything he said” has every instance of “If he said that p , then I believe that p ” as entailments, the former sentence has in its entailment set all and only those instances of this schema which contain “Quine”. Note the intended scope of “only”: it is not that these instances are the only sentences in the entailment set, it is that these are the only instances of the schema that are in the set. No (finite) sentence in a language lacking second order quantification or a synonym to “about” has such a set of entailments, wherefore “about” increases the expressive power of such a language, by (EP). With a second order quantifier, however, we could form the sentence “ $(\forall F)(\text{If he said that } F(\text{Quine}), \text{ then I believe that } F(\text{Quine}))$ ”, which would arguably have the same entailment set as “I believe everything he said about Quine”. The point deflationists usually make here is that the expressions in question achieve the expressive strengthening without any grammatical innovations, such as the introduction of a second order quantifier.

Note that the entailment set of “Everything he says is true” is such that no sentence of the original language has an entailment set that *subsumes* it. This is not so for “I believe everything he said about Quine”, whose entailment set is subsumed by that of “I believe everything he said”. Rather, “about” enables sentences with more *restricted* entailment sets than any original sentence. That this is useful is obvious. If *a* truly asserted that $F_1(b)$ and that $F_2(b)$ and ... and that $F_{1,000,000}(b)$ and falsely asserted that $F_1(c)$ and that $F_2(c)$ and ... and that $F_{1,000,000}(c)$, then it would be nice to be able to have a sentence with only the right instances of “If *a* asserted that *p*, then *p*” as entailments, which is what “about” enables.

What are the inferential properties of the type of sentence recurring above, “*A* φ something about *a*” (where “ φ ” stands proxy for any verb that takes “that”-clauses as objects)? Its entailment set will be rather poor, since it is an existential quantification. Some might want to say that it contains an infinite disjunction of sentences of the form “*A* φ that $S(a)$ ”, but this seems less than clarifying as to the use of “about”, since humans cannot operate with infinite sentences. More relevant are the facts that it is entailed by every sentence of the form “*A* φ that $S(a)$ ”, and contradicts every sentence of the form “*A* only φ that *p*”, which do not contain “*a*”.

Horwich (1998a: 121ff.) gives a different account of the function of “about”, which, however, I find somewhat dubious. In his view, the point of “about” is shown by considering a speaker, *A*, who asserts the sentence “*a* is *F*”. Now suppose that *B* wants to report to *C* what belief of hers that *A* hereby expressed, and that *C* does not understand “*a*”, but does understand the co-referring term “*b*”. In this case, Horwich says, it might be confusing for *C* to be told, “*A* believes that *a* is *F* and $a=b$ ”, since he does not understand “*a*”. Instead, then, *B* can say, “*A* believes about *b* that it is *F*”.

However, if “*a*” is a name, then it does not seem potentially confusing at all to use “*A* believes that *a* is *F* and $a=b$ ” – we introduce new names into others’ vocabulary all the time

simply by *using* them. This is especially unproblematic for the present example, since it is here introduced by an identity sentence. If, on the other hand, “*a*” is a definite description, it is not clear that B would make a true report about her belief. If what A uttered was “The shortest spy is short”, then it is not certain that what is expressed is a belief about someone that he is short. The only case in which the use of “about” would serve this purpose thus seems to be when the attributee expresses a *de re* belief by using a definite description that the hearer does not understand. This is not a persuasive account of the usefulness of “about”.

Horwich’s notion of reference, as defined by his propositional theory, could probably be shown to increase the expressive power of a language in a way matching the story I developed above. But as an account of the usefulness of “about” in natural languages, this account seems implausible given how it requires speakers to operate with the notion of a propositional constituent and the angle brackets that Horwich uses to refer to them.

5. Broader issues

In this section, I will show how the deflationary theory of reference relates to certain well-known issues that are typically taken to turn on the elucidation of reference. I think the situation much resembles that involving truth deflationism. Horwich has argued that his minimal theory remains neutral with respect to many metaphysical issues and concluded that they do not really concern the nature of truth at all. It is not entirely non-committal, however. For instance, a minimalist about truth and falsity must endorse the law of bivalence for propositions, on pain of contradiction, and from this (and some plausible assumptions) follows the law of excluded middle. Thus, minimalism excludes anti-realism in one sense of the word. That a deflationary theory has a contentious consequence, however, does not in itself count against it, as, e.g., Engel (2002) assumes. It would be a strange to take one’s theory as by definition uncontroversial, and no deflationist does. This misconception seems to

derive from the habit among many writers of characterising deflationism as saying that truth is “uninteresting”, “flat”, thin”, etc. Another important point is that it is not really the deflationary part of the theory that entails bivalence – that is, it is not the claim that the truth schema is explanatorily exhaustive, but the schema itself, that entails this. The same holds for all the entailments of my deflationary theory of reference and aboutness.

On the other hand, minimalism is arguably neutral with respect to realism understood, e.g., as the claim that whether something is true or false is dependent on x (mind, language, culture, or whatnot), as long as this does not entail that some propositions are neither true nor false. For to say that whether the proposition that snow is white is true depends on x is just to say that whether snow is white depends on x , but minimalism may well be neutral with respect to both of these dependencies. On the deflationist view, the truth-predicate is used in this connection merely to serve to generalise over sentences of the form “Whether p depends on x ”.

I believe the above deflationary theory of reference is likewise neutral on some issues and not on others. It seems, however, that to the extent that it is not entirely neutral on an issue, it still falls short of deciding the issue, i.e., even with the complete theory of reference and aboutness at hand, we do not know how to settle it. But some issues – those on which the deflationary theory is “not entirely neutral” – must be somewhat reconceived in light of the theory, examples of which will appear below. The reason that most philosophers believe that many important issues turn on the theory of reference is that these notions are often used to express the various standpoints, and often inevitably so. But from the deflationary perspective, this should be explained by appeal to the expressive, or generalising, properties of “about” and (derivatively) “refer”.

Consider “I believe that Pegasus is a winged horse”, which seems true although it contains an empty name. By (A), it follows from this sentence that I believe something about

Pegasus. While this *sounds* right, most would deny what seems to follow from this, namely, that there is something such that I have a certain belief about it. This also has some relevance to the debate between realists and nominalists, for (A) entails that if someone believes that $1+1=2$, then she believes something about 1. But I do not think this should deter nominalists, since our ordinary language is already known to involve much (surface) quantification over *abstracta*.

(A) itself is neutral on how to conceive of the position following “about” in the instances of (A), so we *could* deny the validity of the inference to “There is something such that I have a certain belief about it”. It is not neutral, however, on the hypothetical claim that if I believe that Pegasus is a winged horse, then I believe something about Pegasus. Perhaps this looks like a contentious consequence. However, given that all normal speakers would agree with it, I believe the potential opposite reaction from philosophers derives from a change of use due to new terminological conventions. If we do want to bite the bullet and deny that I have a belief about Pegasus, the deflationary theory of reference would require, firstly, that we also deny that I believe that Pegasus is a winged horse, and secondly, a concomitant explanation of why we found it intuitive. We thus have (at least) two options, neither of which is entirely unreasonable, and both are open to deflationists about reference. Nevertheless, the theory does not allow just any solution to the puzzle.

Another example that shows how an issue may need to be somewhat reconceived is the question of what “water” referred to in the 17th century. To avoid prejudging the issue, we should leave it open whether it referred to water, since “water” in our vocabulary may be different from that of 17th century speakers of English. (For instance, “water”, as used by speakers with a basic knowledge of chemistry, might be interpreted so as to make “water consists of H₂O molecules” come out as an *a priori* necessity, due to a redefinition motivated by certain empirical findings.) Let us use “swater” for the stuff that fills our lakes and taps,

i.e., H_2O , “twater” for the stuff on twin Earth, i.e., XYZ, and “dwater” for all substances sharing the observational features of swater and twater.

Externalists claim that “water” referred to swater. On the above theory of reference, firstly, this is to be understood as meaning that they used ‘water’ to say things about (i.e., refer to) swater. But this use of “about” is merely a device of generalisation, namely, over the infinite conjunction “They used ‘water’ to say that swater is wet and that swater is cold and ... (and so on *ad infinitum*)”. Thus, the issue is not primarily about aboutness but about what they said when using “water”. Whether “water” was used to refer to H_2O will then depend on whether it was used to refer to swater, whether swater = H_2O , and whether the position after “about” is extensional. (I use “swater” rather than “ H_2O ” in order to avoid certain questions about the semantics of attitude ascriptions and the *de re-de dicto* distinction. Perhaps 17th century speakers believed that swater was wet but not that H_2O was wet.)

Given the deflationary theory of reference presented above, even an exhaustive theory about reference does not settle whether the externalist is right. In that way, then, it is neutral. On the other hand, the theory entails that the issue must be conceived of in a particular way, which, in turn, might make one answer more plausible than others.

However, a different consideration seems to suggest that accepting the above theory of reference actually only changes the terminology of the debate. I said above that this theory does not force us to shun all questions about the relations between expressions and external objects. It only denies that reference is such a relation. But we are of course free to ask in virtue of what a name can be used to say certain things. We are not denying that “Aristotle” can be used to say that Aristotle is a philosopher, and that Aristotle is Greek, and so on; in other words, to say things about Aristotle. Perhaps, then, there some relation between “Aristotle” and Aristotle in virtue of which the former can be used to say things about the latter, and which can be given an informative analysis. Otherwise put, there might be a

relation between “Aristotle” and Aristotle the description of which would explain why “Aristotle is Greek” is used to say that Aristotle is Greek rather than that Plato is Greek (and so on). Let us call this relation *denotation*. Could it be that the analysis of denotation, if such can be had, will tell us what “water” was used to refer to?

Maybe, but the issue might also be settled otherwise. The question of what beliefs our ancestors had seems independent of the nature of denotation. Did they (mainly) believe things about swater or about dwater? If it can somehow be shown that they had no beliefs about swater, but only about dwater, then it is very likely that they used “water” to refer to dwater and not swater. In this case, not only is the elucidation of reference or aboutness irrelevant for answering the original question, but also the elucidation of denotation. (However, if there is a true, informative analysis of denotation, we might be guided to it by the finding that their beliefs were about dwater rather than swater).

But we should not conclude from this that the theory of denotation *must* take claims about what beliefs they had as data that must be settled in advance. For perhaps a certain hypothesis about denotation gives a very plausible account of certain uncontroversial facts, e.g., that “Aristotle” can be used to refer to Aristotle. Suppose also that this hypothesis, if applied to “water”, as used in the 17th century, entails that “water” denotes dwater, rather than swater. If we would accept this claim on the grounds that the uncontroversial facts were so well accommodated, we would have to accept that “water” was used to refer to dwater, and then, presumably, it would be reasonable to conclude that they had beliefs only about swater. Again, we have an illustration of standpoints on a certain issue that the deflationary theory of reference allows, but among which it does not choose.

These issues are about reference only in the sense that one can generalise over what is centrally at issue by using “refer” or “about”. Further, while the theory entails that the *nature* of reference or aboutness is not at issue in these debates, this need not change the

philosophical landscape as much as it may first seem. For it may well be that the nature of *denotation*, as defined above, is crucial.

Much of the above presupposes for the sake of argument that there is a correct, informative analysis of denotation. But this is of course something on which deflationism about reference is neutral. Even if it turns out that there is no informative reduction of denotation, however, there may still be other, intimately related issues about how we relate mentally and linguistically to external objects, and how these relations might be analysed, which, similarly, are not prejudged by the deflationary theory of reference. The notion of *de re* belief is often explained using the word “of”, where this is used as a synonym to “about”, as in the phrase, “*a* believes *of b* that ...it...”. But *de re* beliefs may equally, and perhaps less misleadingly, be identified without the use of “about” or synonymous expressions. For instance, instead of saying “*a* believes *of b* that ...it...”, one can say, “*b* is such that *a* believes that ...it...”, or, “ $\exists x(x = b \ \& \ a \text{ believes that } \dots x \dots)$ ”.

The phenomenon of *de re* belief so understood seems to concern the scope of “believes” rather than the word “of” or “about”. But, as can be seen by examining our intuitions about *de re* beliefs in this sense, these issues are still centrally about how we relate mentally to external objects. Now, the deflationary theory of reference does not entail that we cannot give informative necessary and sufficient conditions for there being something such that one believes that it is so and so, or that we cannot explain the difference between believing that something is such and so, and, on the other hand, there being something that one believes to be thus.

Similarly, one can speak of *de re* assertions, as reported by sentences of the form “There is something such that *a* said that it is *F*”, which may also well be analysable, as far as the deflationary theory of reference is concerned. Its only consequence is that this question does not pertain in any way to the notions of aboutness or reference, except that the words

expressing them can be used to form answers to it that would otherwise be infinitely long. Thus, even in the absence of a reduction of denotation, the deflationary theory of reference does not confine us to deflationary answers, or to a Tractarian silence, on all questions concerning how we relate, mentally and linguistically, to the external world.^{vi}

ⁱ Admittedly, it is more controversial that the equivalence between (3a) and (3b) should follow. The motivation is that sentences equivalent in virtue of meaning are synonymous, wherefore, by compositionality, they ought to be intersubstitutable in all (but quotational) contexts. Many, of course, deny that synonymous expressions can be thus intersubstitutable. However, this idea is based on Benson Mates's (1952) consideration that, intuitively, one could believe that Greeks are Greeks without believing that Greeks are Hellenes, etc. As argued in Båve (2008), however, this intuition is likely to have a pragmatic explanation. Since we have good reason to believe in a principle of compositionality that entails the intersubstitutivity of synonyms, we should thus regard the consequence of disquotationalism as a weakness. (Readers who doubt this substitutivity should note that this is but one of many arguments against truth-disquotationalism.) Further, if we take instead "That *p* is true" to be the primary form of truth-ascription, then a good case can be made that the counterparts of (3a) and (3b) are equivalent (as concerns the intuition that one can believe that snow is white without believing that it is *true* that snow is white, see Båve (2006: 131f.). If they are indeed intersubstitutable, then the disquotationalist owes us an answer to why this is not the case on his theory (an answer not likely to be forthcoming). If they are intersubstitutable, further, that is in itself a reason to believe that truth-ascriptions to propositions are primary (and, so, this truth-theory is *ceteris paribus* to be preferred).

ⁱⁱ If such a propositional deflationary theory is also formulated purely linguistically, e.g., by the claim that the instances of the truth-schema are analytic, or some such, then it does not

even come with a commitment to propositions, avoiding which is the major motivation for disquotationalism (see Båve (2006: Ch. 5)).

ⁱⁱⁱ If there is a worry about the use of “refer” in this sentence, let us instead put it this way: there should be an expression replacing “*x*” which is a *hyponym* to “propositional constituent”. That is, supposing the term in question is “*t*”, and given a definition of “propositional constituent”, the sentence “*t* is a propositional constituent” should come out as assertible/analytic.

^{iv} The use of “singular term” may seem problematic. Firstly, this notion cannot be understood in terms of reference, because the theory of “refer” to be given below defines it in terms of “about”, so if “singular term” is understood in terms of reference, the combination of the two theories would be viciously, or at least unnecessarily, circular. Instead, the notion could be defined inferentially or intralinguistically, e.g., as in Frege (1884: §§ 26, 60-62), Dummett (1973: 57ff.), Geach (1975), Wright (1983: ix), Hale (1987: Ch. 2), Hale and Wright (2001: Chs. 1 and 2), Brandom (1994: Ch. 6), or Heck (unpublished).

A second problem concerns definite descriptions. One option is of course to deny that definite descriptions are singular terms. This could be motivated by the conjunction of (1) Russell’s analysis, on which “The *F* is *G*” means the same as “There is exactly one *F*, which is *G*” (or perhaps the view that they are existential quantifiers) and, (2) the view that the proposition that there is exactly one *F* that is *G* is not *about* anything (since it is purely quantificational). Then, definite descriptions would be no counter-example to (A).

If, however, we follow the majority in calling them singular terms, then one instance of (A) would be “That the shortest spy is a spy is about the shortest spy”. Whether this is a problem, however, depends on one’s further views. One could say that definite descriptions, perhaps contrary to appearances, are “Millian”, i.e., that they contribute to which proposition is expressed by sentences containing them merely by providing an object “the one satisfying

the description” (and that these propositions are *singular* with respect to that object). In that case, Russell’s analysis is false and we ought to have no qualms about accepting the instances of (A) containing definite descriptions.

A third option is to agree with Russell’s analysis and the view that definite descriptions are singular terms, and simply accept that the proposition that there is exactly one F, which is G, is *about* the F, after all. Given that Russell’s analysis and the view that definite descriptions are singular term really are true, this does not seem all that outrageous. It may be objected, however, that on the deflationary theory of aboutness, we could then not explain *how* the proposition that there is exactly one F which is G can be *about* the F. But I think we can. First, we can assume the (A)-instance “That the F is G is about the F” (since descriptions are assumed to be singular terms), and then, we substitute in this instance “The F is G” by “There is exactly one F which is G”, which is valid if the Russellian analysis is true, since it states a *semantic* (rather than, say merely a material) equivalence.

A fourth option is to agree that definite descriptions are singular terms, but argue that this category subdivides into “A-terms” (that *do* validate (A)) and “B-terms” (that *don’t*), and that “t” in (A) stands proxy for the A-terms. The A-terms would then include names, indexicals, and perhaps more, and the B-terms would include definite descriptions, and perhaps more. Of course, we could not *define* “A-term” as an expression that validates (A), since (A) would then become trivial. Presumably, however, there is some semantic (perhaps inferential) feature of definite descriptions in virtue of which they do not validate (A), and then, one could define the A-terms in terms of that feature. (Of course, we might as well define “singular term” that way and stick to the original formulation, which would only be a notational variant of the theory speaking of A-terms.) That the wider category of expressions (under whichever name) have some fundamentally different semantic functioning is granted by the very admission that they do not all validate (A).

There is some uncertainty, then, over what the instances of (A) should be. But I do not think this is as bad as it seems. First, none of the four possibilities I have discerned seem to go against the spirit of the deflationary theory of aboutness, wherefore definite descriptions, as far as we can tell, will not pose a problem for it. Secondly, the reason we do not know which expressions instantiate (A) is that the question depends on the semantic functioning of definite descriptions, an issue that the theory of reference should not prejudge.

^v This take on reference was anticipated by Austin, Strawson, and, in particular, Searle, who wrote, “To say that an expression refers (predicates, asserts, etc.) in my terminology is either senseless or is shorthand [*sic!*] for saying that the expression is used by speakers to refer” (1969: 28). I would not add the qualification “in my terminology”, however, since I believe this simply is *the* way “refer” is understood.

None of this, however, is to deny that there may be some interesting relation between an expression and an object in virtue of which speakers use it to refer to the object (or some other relation philosophers may have had in mind when using “reference”). That is, I am not taking a stand in the debate between the abovementioned philosophers and defenders of technical notions of reference. I do claim that Strawson *et al.* were right about ordinary usage, but I don’t think anyone disagreed with *that* claim, only with the way he used it to criticise certain semanticists and truth-theorists. Since I am engaged in natural language semantics, I am obliged to respect the data, i.e., ordinary speakers’ use and intuitions. But this is not to say that it is unacceptable to define a notion for technical/scientific purposes. More on this in section 5.

^{vi} Thanks to two anonymous referees at *Synthèse* for valuable comments on earlier drafts.

REFERENCES

- Båve, A. (2006), *Deflationism: A Use-Theoretic Analysis of the Truth-Predicate*, Ph. D. Thesis, Acta Universitatis Stockholmiensis, Stockholm, Almqvist & Wiksell International.
- Båve, A. (2008), “A Pragmatic Defense of Millianism”, *Philosophical Studies* **138**, pp. 271-289.
- Båve, A. (forthcoming), “Why Is A Truth-Predicate Like A Pronoun?”, forthcoming in *Philosophical Studies*.
- Brandom, R. (1994), *Making It Explicit*, Cambridge, MA, Harvard University Press.
- (2002), “Explanatory vs. Expressive Deflationism about Truth”, in Schantz (2002), pp. 103-19.
- David, M. (1994), *Correspondence and Disquotation*, Oxford, Oxford University Press.
- Dummett, M. (1973), *Frege: The Philosophy of Language, Second Edition*, Cambridge, MA., Harvard University Press.
- Dunn, J. M. and A. Gupta (eds.), *Truth or consequences*, Dordrecht, Kluwer Academic Publishers.
- Engel, P. (2002), *Truth*, Montreal, McGill-Queen’s University Press.
- Field, H. (1994), “Deflationist Views about Meaning and Content”, *Mind* **103**, pp. 249-85.
- Frege, G. (1884/1950), *The Foundations of Arithmetic*, trans. by J. L. Austin, Cambridge, Basil Blackwell.
- Geach, P. (1975), “Names and Identity”, in S. Guttenplan (ed.), *Mind and Language: Wolfson College Lectures 1974*, Oxford, Clarendon Press.
- Grover, D., Camp, J. and Belnap, N. (1975), “A Prosentential Theory of Truth”, *Philosophical Studies* **27**, pp. 73-125.
- Gupta, A. (1993), “A Critique of Deflationism”, *Philosophical Topics* **21**, pp. 57-81.
- Hale, B. (1987), *Abstract Objects*, Oxford, Basil Blackwell.

- Hale, B. and C. Wright (eds.) (2001), *The Reason's Proper Study: Essays towards a Neo-Fregean Philosophy of Mathematics*, Oxford, Oxford University Press.
- Heck, R. (unpublished), "What Is A Singular Term?", unpublished manuscript.
- Horwich, P. (1998a), *Meaning*, Oxford, Clarendon Press.
- (1998b), *Truth*, second edition, Oxford, Clarendon Press.
- (2005), *Reflections on meaning*, Oxford, Clarendon Press
- Kirkham, R. (1992), *Theories of Truth*, Cambridge, MA., MIT Press.
- Leeds, S. (1978), "Theories of Reference and Truth", *Erkenntnis* **13**, pp. 111-30.
- Linsky, L. (ed.) (1952), *Semantics and the philosophy of language*, Urbana, University of Illinois Press.
- Mates, B. (1952), "Synonymity" in Linsky (1952), pp. 111-136.
- Schantz, R. (ed.) (2002), *What is Truth?*, Berlin, de Gruyter.
- Searle, J. (1969), *Speech Acts*, Cambridge, Cambridge University Press.
- Wilson, W. K. (1990), "Some Reflections on the Prosentential Theory of Truth", in Dunn and Gupta (1990), pp. 19-32.
- Wright, C. (1983), *Frege's Conception of Numbers as Objects*, Aberdeen, Aberdeen University Press.