

Translatable Self-Reference

HARTLEY SLATER

UNIVERSITY OF WESTERN AUSTRALIA

hartley.slater@uwa.edu.au

Received by Greg Restall

Published May 19, 2011

<http://www.philosophy.unimelb.edu.au/ajl/2011>

© 2011 Hartley Slater

Abstract: Stephen Read has advanced a solution of certain semantic paradoxes recently, based on the work of Thomas Bradwardine. One consequence of this approach, however, is that if Socrates utters only ‘Socrates utters a falsehood’ (a), while Plato says ‘Socrates utters a falsehood’ (b), then, for Bradwardine two different propositions are involved on account of (a) being self-referential, while (b) is not. Problems with this consequence are first discussed before a closely related analysis is provided that escapes it. Moreover, this alternative analysis merely relies on quantification theory at the propositional level, so there is very little to question about it. The paper is the third in a series explaining the superior virtues of a referential form of propositional quantification.

Stephen Read has advanced a solution of certain semantic paradoxes recently, based on the work of Thomas Bradwardine [4, 5, 6, 7]. Read claims that Tarski’s T-scheme, in the form

$$(T) \quad T(s) \equiv A$$

must be improved so that, while Tarski “said that (T) should hold when what replaces ‘A’ is a translation (into the metalanguage) of the sentence whose name is ‘s’ . . . Bradwardine’s point is that what replaces ‘A’ in (T) must express everything that the utterance named by ‘s’ expresses or signifies” [7, p. 166]. Writing ‘Sig(s, A)’ to express the fact that sentence (‘utterance-type’) *s* signifies that *A*, Read defines the truth and falsity of sentences in the following fashion:

$$\text{Tr}(s) \equiv (\exists A)\text{Sig}(s, A) \ \& \ (\forall A)(\text{Sig}(s, A) \supset A)$$

$$\text{Fa}(s) \equiv (\exists A)(\text{Sig}(s, A) \ \& \ \neg A)$$

He then takes Bradwardine’s major principle to mean that if *A* entails *B* then Sig(*s*, *A*) entails Sig(*s*, *B*), so that while, for instance, in a paradoxical case, Tarski would think that a given sentence meant just that it was false, or not true, Bradwardine would start from the assumption that the sentence might

express or signify not only that but something else, as well: $\text{Sig}(s, \text{Fa}(s) \ \& \ Q)$. In fact it follows from Bradwardine's principles that no sentence can signify just that that it is false, or not true. After resolving an impressive number of puzzles on this basis, Read has summed up: 'Once truth and meaning are properly understood, so that an utterance is true only if everything it signifies obtains, paradox is prevented without need for logical revision' [7, pp. 175–6].

One consequence of this approach, however, is that if Socrates utters only 'Socrates utters a falsehood' (a), while Plato says 'Socrates utters a falsehood' (b), then, for Bradwardine two different propositions are involved on account of (a) being self-referential, while (b) is not. It cannot be the same proposition that is involved since, on Bradwardine's account, the first is false while the second is true. Thus Read says ([7, p. 171]—though take care, the two truth-values of (a) and (b) are accidentally reversed on line 3 of this page):

Bradwardine's response was that (a) and (b) were not the same proposition, since (a) is self-referential, saying of itself that it is false, and so, by Thesis (2), saying also that it is true. That is why it is false. But (b) is not self-referential, but says (only) that (a) is false (and uttered by Socrates). . . . Superficially, (a) and (b) are tokens of the same utterance type, so it might seem that we have here a counterexample to my earlier claim that Bradwardine attributed truth and signification to utterance types. However, we have just seen that (a) and (b) differ in signification. So they are tokens of different sentence-types.

But what is another token (if there is one) of the sentence-type that Socrates is then supposed to utter? Since (a) is a token sentence, Socrates' uttering the same words again would, on Read's analysis, involve self-reference to another, different token, and so could not have the same signification as (a). What does have the same signification, supposedly, is '(a) is false and (a) is true and (a) is uttered by Socrates'. But that is certainly not of the same sentence-type as (a).

And does it even have the same content? That is the more important question. Read seems, in fact, to be forgetting the rules about translation into reported speech, and specifically confusing a proper name with an egocentric particular. One must take the utterer into account to get the content of what is said if the speaker says 'I am uttering a falsehood'. But nothing needs to be changed if a proper name is involved, as in 'Socrates utters a falsehood'. How else is anyone going to report *what Socrates said* in the given case if not by uttering what Plato did, or its equivalent?

Maybe Read, after a re-consideration, would like to say that what Socrates says in the case above is in fact untranslatable. This option is taken by 'cassationists' [2], and would mean saying that Socrates' utterance did not express a proposition, making Plato's remark false, since Socrates' was neither true nor false. But surely Socrates *agrees* with Plato; he can even say 'Yes!' or 'That's true' after Plato's utterance. So Socrates must be able to have the same thought as

Plato; indeed, as before, the rules of reported speech require that in the context above both have the same thought. So it has to be pointed out that there is another treatment of the same puzzles Read has been dealing with that is much more in Bradwardine's spirit, yet does not get into the same difficulties [8, 11].

I use an improved form of Prior's propositional quantification incorporating Cocchiarella's nominalising functor ' λ ' ('that', see [1, pp. 216–7]), which makes ' λp ' a phrase referring to the proposition expressed by ' p ' on the occasion of its use, with the standard interpretation. Propositional identity statements of the form ' $\lambda p = \lambda q$ ' are true only if ' p ' and ' q ' are synonymous, which involves a tighter relation than logical equivalence. But there are many other propositional identity statements that are not of this form, since there are many other ways to refer to propositions than by using 'that'-clauses. Here I also use the epsilon calculus (rather than the predicate calculus) to provide other propositional referential terms; together these two improvements enable one to generate all the identity statements that are needed. As we shall see, it is the provision of propositional referential terms that chiefly separates this alternative treatment from Read's, and specifically in the above and similar cases it allows the extra significance of the puzzling utterances to be pointed towards while leaving it un-expressed, thereby capturing in a more appropriate way its 'non-translatability'. I write ' Sxr ' for '(sentence) x signifies (proposition) r ', so that ' $\epsilon r Sxr$ ' means 'what x signifies' (which might be a matter of choice if there are several things x signifies). General equivalences are then that $\top \lambda p \equiv p$, and that $Sx\epsilon r Sxr \equiv (\exists r) Sxr$.

Read does not use propositional referential terms, as we saw at the start; nor does he have a truth predicate with propositions. Thus his analysis involves quantification over expressed sentences, as with ' $(\exists A) \text{Sig}(s, A)$ ', in the manner of Prior, and the truth of propositions is not represented as a predicate, but simply by the assertion of the proposition, if at all. That leads him to bring in what extra might be involved in a case of self-reference by means of some further expressed sentence: viz ' $\text{Sig}(s, Fa(s) \ \& \ Q)$ '. From this, using the principles indicated above, Read easily draws the conclusion $Fa(s)$ [7, p. 169], leading him to take this to be the case when Socrates utters 'Socrates utters a falsehood', but, of course, not to be the case when Plato utters the same thing. So he is drawn to analyse Socrates' remark as though it was untranslatable, in the manner we saw before.

A couple of elementary results, for a start, will show how the closely related approach using propositional referential terms does not have this unfortunate consequence. First it is true that if a sentence x signifies that what that sentence signifies is false (or not true) then there is not just one thing the sentence signifies, but one thing that is true and another thing that is false. This would arise, for example, with 'What this sentence signifies is not true', on a self-referential reading of 'this sentence'. For if $Sx\lambda(\exists r)[Sxr \ \& \ (t)(Sxt \supset t = r) \ \& \ \neg Tr]$, then, if $(r)(Sxr \supset Tr)$, then, by substitution, $\top \lambda(\exists r)[Sxr \ \& \ (t)(Sxt \supset t = r) \ \& \ \neg Tr]$, and so in particular $(\exists r)(Sxr \ \& \ \neg Tr)$, a contra-

diction. So $\neg(r)(Sxr \supset Tr)$, i.e. $(\exists r)(Sxr \ \& \ \neg Tr)$. But the latter, together with $Sx\lambda(\exists r)[Sxr \ \& \ \neg Tr]$, means that $(\exists r)(Sxr \ \& \ Tr)$. But that means that $(\exists t)(\exists r)(Sxt \ \& \ Sxr \ \& \ \neg Tt \ \& \ Tr)$, and so $(\exists t)(\exists r)(Sxt \ \& \ Sxr \ \& \ t \neq r)$.

Of course, from $(\exists r)(Sxr \ \& \ \neg Tr)$ one can derive $Sx\epsilon r(Sxr \ \& \ \neg Tr) \ \& \ \neg T\epsilon r(Sxr \ \& \ \neg Tr)$, while from $(\exists r)(Sxr \ \& \ Tr)$ one can get $Sx\epsilon r(Sxr \ \& \ Tr) \ \& \ T\epsilon r(Sxr \ \& \ Tr)$, so identifying the leading candidates for $\epsilon r(Sxr \ \& \ \neg Tr)$ and $\epsilon r(Sxr \ \& \ Tr)$ is quite straightforward since $Sx\lambda(\exists r)[Sxr \ \& \ (t)(Sxt \ \supset \ t = r)]$, while $\neg T\lambda(\exists r)[Sxr \ \& \ (t)(Sxt \ \supset \ t = r)]$, and $Sx\lambda(\exists r)[Sxr \ \& \ \neg Tr]$, while $T\lambda(\exists r)[Sxr \ \& \ \neg Tr]$.

Notice also that parallel things would arise given $Sy\lambda(\exists r)[Sxr \ \& \ (t)(Sxt \ \supset \ t = r) \ \& \ \neg Tr]$, i.e. with, for example, ‘What that sentence signifies is not true’ with the previous sentence as the referent of ‘that sentence’. For $Sy\lambda(\exists r)[Sxr \ \& \ (t)(Sxt \ \supset \ t = r)]$, and $Sy\lambda(\exists r)[Sxr \ \& \ \neg Tr]$, and what y says here are, respectively, false and true. So in this case the report given by y carries completely the same message as the original expression by x , i.e. the content of x is unproblematically translatable.

A different result is obtainable supposing merely that a sentence signifies that something it signifies is not true, i.e. $Sx\lambda(\exists r)(Sxr \ \& \ \neg Tr)$, which is the same case as that above without the uniqueness clause. Certainly then, again, if $(r)(Sxr \ \supset \ Tr)$, then, by substitution, $T\lambda(\exists r)(Sxr \ \& \ \neg Tr)$, and so $(\exists r)(Sxr \ \& \ \neg Tr)$, a contradiction. Hence $\neg(r)(Sxr \ \supset \ Tr)$, i.e. $(\exists r)(Sxr \ \& \ \neg Tr)$. And the latter, together with what was explicitly said in this case, means that $(\exists r)(Sxr \ \& \ Tr)$. So again something x signifies is false, and something x signifies is true. But while, again, one thing that is said that is true is the content of the sentence, i.e. $\lambda(\exists r)(Sxr \ \& \ \neg Tr)$, whatever is said that is false is not immediately expressible in a ‘that’-clause, but instead is just referred to by the epsilon term in ‘ $Sx\epsilon r(Sxr \ \& \ \neg Tr) \ \& \ \neg T\epsilon r(Sxr \ \& \ \neg Tr)$ ’. And now no entirely comparable result is entailed if $Sy\lambda(\exists r)(Sxr \ \& \ \neg Tr)$, since then, while it follows that $(\exists r)(Syr \ \& \ Tr)$, there is no way to obtain that $(\exists r)(Syr \ \& \ \neg Tr)$.

Turning now to Read’s case above, we find it is the same as this last one, since it is about a sentence that says not about what it says, but about itself that it is false. Writing $Vs \equiv (r)(Ssr \ \supset \ Tr)$, then given $Ss\lambda\neg Vs$ we immediately get from the last case that $(\exists r)(Ssr \ \& \ \neg Tr)$, and that $(\exists r)(Ssr \ \& \ Tr)$. The former result shows that in fact $\neg Vs$ (Read’s ‘ $Fa(s)$ ’), and it is from this fact, with $Ss\lambda\neg Vs$, that we get that there is something Socrates says that is true, namely $\lambda(\exists r)(Ssr \ \& \ \neg Tr)$. Given $Sp\lambda\neg Vs$, that means also that $(\exists r)(Spr \ \& \ Tr)$, for the same substitution instance, showing Socrates and Plato do then say one and the same thing that is true. But what it is that Socrates implies that is false in uttering what he did is not so easily locatable, and in fact what anyone might add to get something equivalent to the whole of what Socrates is implying is not obtainable from the given linguistic context, since it is not expressed explicitly there.

It is at this point that the present account wins out over Read’s, since there is still something Plato might add to express an equivalent to what Socrates

is thinking—if only Socrates would make it explicit in a ‘that’-clause. So what Socrates is thinking as a whole is not untranslatable; it is merely not wholly translatable given the limited description of the context presented above. The central formal difference from Read’s account which allows for this is that the limited context above, even if it does not give us what proposition is implied that is false, still does give us *a means of referring to that proposition*, and also ensures us that that proposition is a different proposition from the one explicitly made. For $\neg\exists r(Ssr \ \& \ \neg Tr)$, while $\exists r(\lambda(\exists r)(Ssr \ \& \ \neg Tr))$. But only the speaker, Socrates, can help us to locate $\epsilon r(Ssr \ \& \ \neg Tr)$ explicitly, and so express it, since by just uttering ‘Socrates utters a falsehood’ he has yet to spell out which falsehood is on his mind. In this connection it is important to note that while the remark is ‘self-referential’ in that it refers to the speaker making it, and also to the sentence that is used to express it, there is no indication that what proposition is taken to be false is what is then said, i.e. that $Ss\lambda(y)(Ssy \ \supset \ y = \lambda(\exists r)(Ssr \ \& \ \neg Tr))$. If that had been added then we would have a case like the initial one analysed, where there was an explicit mention of something said that was false. But the case Read has described is strictly without this further form of ‘self-reference’, making what is implied that is false indeterminate, from what has so far been given.

Quite generally, if a speaker makes an existential remark like ‘ $(\exists r)Pr$ ’, then we can ask ‘Which P is it?’ with the ‘it’ then in question being tracked by the associated epsilon term ‘ ϵrPr ’; and if a speaker says ‘ $\neg(r)Pr$ ’ then we can ask ‘Which is it that is not P?’ with the ‘it’ in question being tracked by ‘ $\epsilon r\neg Pr$ ’. It is something of the former kind which is a candidate for being the extra, unexpressed thought Socrates has that is not true when he utters ‘ $(\exists r)(Ssr \ \& \ \neg Tr)$ ’. *That Socrates was thinking something that is false* is demonstrably true, but he needs to tell us *what he was thinking that he was thinking and was false*, before we have a chance of telling him explicitly what he was thinking that is false. There are two cases, the first being the case above where uniqueness is a part of the very thought he is then actually having, and is just what is false. That is to say we have $Ss\lambda(Sst \ \& \ \neg Tt)$ with $t = \lambda(\exists r)(Ssr \ \& \ \neg Tr)$, while $\exists r(\lambda(\exists r)(Ssr \ \& \ \neg Tr))$. But it also may be that he thinks he is implying something else by the remark we have been considering, and thinks that that proposition is false, i.e. that $Ss\lambda(Sst \ \& \ \neg Tt)$ for a quite separate $t \neq \lambda(\exists r)(Ssr \ \& \ \neg Tr)$. Since, however, he is not implying anything else, this implication is imaginary and we have $\neg(Sst \ \& \ \neg Tt)$, so ‘ $\lambda(Sst \ \& \ \neg Tt)$ ’ provides the needed truth maker for ‘ $(\exists t)(Ssr \ \& \ \neg Tr)$ ’.

One needs to be familiar with the epsilon calculus to realise fully that making an existential remark always means referring, or at least alluding to an instance, since the predicate calculus hasn’t got an automatic placeholder for what individual thing is being spoken about when someone makes such a remark. Intuitionists do not appreciate this point fully, and in particular would have difficulty with a request for a witness to a negated universal remark in the manner presented before. The epsilon calculus does not sit easily with Intu-

itionistic Logic (see [12, p. 397–402] for a full discussion) and the matter is settled against this logic if it is granted that the previous question ‘Which is it that is not P?’ can be appropriately asked after a negated universal remark. For even if the identity of $\epsilon r \neg Pr$ is unknown, at least it is then granted that $\neg P \epsilon r \neg Pr$, and so $(\exists r) \neg Pr$. But the further problem in Read’s case is that the existential quantifier is within the scope of an intensional operator, so any instantiation which the speaker might produce is up to him to say, and remains a feature of the broader pragmatic context. It may seem that there is nothing similar in the earlier case without a uniqueness clause, since no speaker was nominated, and how could a bare sentence provide an explicit answer to what further it implicitly means in that case? But no sentence on its own is self-referential [9], since one cannot, for instance, derive by objectival Existential Generalisation “ $(\exists z)(z = ‘(\exists r)(S_z r \ \& \ \neg Tr)’)$ ” from “ $x = ‘(\exists r)(S_x r \ \& \ \neg Tr)’$ ”. The presence of the quotes prevents this, but also, thereby, the ‘x’ in quotes might refer to a quite different sentence, as, equally, ‘this sentence’ in ‘what this sentence signifies is not true’ would do if it were accompanied with a gesture to another sentence. So to get the supposed referent of ‘x’ and make it the case (in indirect speech) that $S_x \lambda (\exists r)(S_x r \ \& \ \neg Tr)$, the sentence ‘ $(\exists r)(S_x r \ \& \ \neg Tr)$ ’ must be uttered by someone giving it a self-referential interpretation. Therefore to get an answer to the question about what $\epsilon r(S_x r \ \& \ \neg Tr)$ is, one must ask that speaker.

The need to consider the speaker’s intentions to secure the self-reference is centrally relevant in connection with other cases, for instance when speaker M says ‘Everything thought by M is false’. For in order to ensure that the quoted ‘M’ is used to refer to the speaker in question the speaker must realise that the universal thought expressed here is one of his or her thoughts. Forgetting this we might symbolise the content of the expression by just reporting: $S_m \lambda (r)(S_m r \supset \neg Tr)$. The analysis of this proceeds much as before, since if we suppose that $(r)(S_m r \supset \neg Tr)$ then, by substitution $(\exists r)(S_m r \ \& \ Tr)$, a contradiction, proving that $(\exists r)(S_m r \ \& \ Tr)$. But that, together with the initial report then means that $(\exists r)(S_m r \ \& \ \neg Tr)$. The instance that verifies the latter existence claim is thereby assured, but what is the instance that verifies the former existence claim, i.e. that $(\exists r)(S_m r \ \& \ Tr)$? It is not available from the context as so far formalised, and has to be derived from what the above incomplete report has missed out: $S_m \lambda S_m \lambda (r)(S_m r \supset \neg Tr)$. For that, with what was previously reported entails $(\exists r)(S_m r \ \& \ Tr)$. A similar implication was available in the other case, namely that $S_s \lambda S_s \lambda (\exists r)(S_s r \ \& \ \neg Tr)$, although this was not needed in the analysis above.

To think that some sentences involved in problems to do with self-reference must have a speaker whose unvoiced thoughts are relevant to the case is not commonly realised. But I suppose some people might also have to adjust to the complex, delicate subtlety of the analysis of this sort of situation that is now clearly available formally. It really is quite remarkable, I agree. But it is just a symbolisation of natural language, so it is that natural language that such

people should better respect. For the record, many other paradoxes are dissolved this way, such as Curry's Paradox (which Read deals with elsewhere [7]) and also, for instance, The Knower Paradox, and Gödel's Paradox [10]. Two other major problems for Priorian propositional quantification I have shown to disappear with the more sophisticated, referential version [8, 13]—Read has yet to address either of these issues. The only relevant paradox I have not previously discussed is Field's Paradox, but that is resolved in much the same way as in [7, 175].

REFERENCES

- [1] Cocchiarella, N. 2006, *Logical Investigations of Predication Theory and the Problem of Universals*, Bibliopolis, Naples.
- [2] Goldstein, L. 2006, 'Fibonacci, Yablo and the Cassationist Approach to Paradox,' *Mind* 115, 867–890.
- [3] Prior, A. N. 1971, *Objects of Thought*, O.U.P., Oxford.
- [4] Read, S. 2002, 'The Liar Paradox from John Buridan back to Thomas Bradwardine,' *Vivarium* 40, 189–218.
- [5] Read, S. 2008, 'The Truth Schema and the Liar,' in S. Rahman, T. Tulenheimo and E. Genot (eds) *Unity, Truth and the Liar*, Springer, Berlin.
- [6] Read, S. 2008, 'Further Thoughts on Tarski's T-scheme and the Liar' in S. Rahman, T. Tulenheimo and E. Genot (eds) *Unity, Truth and the Liar*, Springer, Berlin.
- [7] Read, S. 2010, 'Field's paradox and Its Medieval Solution,' *History and Philosophy of Logic* 31, 161–176.
- [8] Slater, B. H. 2001, 'Prior's Analytic Revised,' *Analysis* 61, 86–90.
- [9] Slater, B. H. 2002, 'Syntactic Liars,' *Analysis* 62, 107–9.
- [10] Slater, B. H. , 2004, 'Ramseying Liars,' *Logic and Logical Philosophy* 13, 57–70.
- [11] Slater, B. H. 2008, 'Out of the Liar Tangle,' in S. Rahman, T. Tulenheimo and E. Genot (editors) *Unity, Truth and the Liar*, Springer, Berlin.
- [12] Slater, B. H. 2009, 'Hilbert's Epsilon Calculus and its Successors,' in D. Gabbay and J. Woods (eds) *Handbook of the History of Logic, Vol 5*, Elsevier Science, Burlington MA, 385–448.
- [13] Slater, B. H. 2010, 'Ontological Discriminations,' *LOGICA Yearbook 2009* Czech Institute of Philosophy, Prague.

The *Australasian Journal of Logic* (ISSN 1448-5052) disseminates articles that significantly advance the study of logic, in its mathematical, philosophical or computational guises. The scope of the journal includes all areas of logic, both pure and applied to topics in philosophy, mathematics, computation, linguistics and the other sciences.

Articles appearing in the journal have been carefully and critically refereed under the responsibility of members of the Editorial Board. Only papers judged to be both significant and excellent are accepted for publication.

The journal is freely available at the journal website at

<http://www.philosophy.unimelb.edu.au/ajl/>.

All issues of the journal are archived electronically at the journal website.

SUBSCRIPTIONS Individuals may subscribe to the journal by sending an email, including a full name, an institutional affiliation and an email address to the managing editor at ajl-editors@unimelb.edu.au. Subscribers will receive email abstracts of accepted papers to an address of their choice. For institutional subscription, please email the managing editor at ajl-editors@unimelb.edu.au.

Complete published papers may be downloaded at the journal's website at <http://www.philosophy.unimelb.edu.au/ajl/>. The journal currently publishes in pdf format.

SUBMISSION The journal accepts submissions of papers electronically. To submit an article for publication, send the \LaTeX source of a submission to a member of the editorial board. For a current list of the editorial board, consult the website.

The copyright of each article remains with the author or authors of that article.