HISTORY AND PHILOSOPHY OF LOGIC, **25** (MAY 2004), 95–110

Taylor & Francis
Taylor & Francis Group

Consequences of a closed, token-based semantics: the case of John Buridan[1]

GYULA KLIMA
Department of Philosophy, Fordham University, 441 E. Fordham Road, New York, NY 10458, USA

Received 12 May 2003 Revised 25 September 2003

This paper argues for two principal conclusions about natural language semantics based on John Buridan's considerations concerning the notion of formal consequence, that is, formally valid inference. (1) Natural languages are essentially semantically closed, yet they do not have to be on that account inconsistent. (2) Natural language semantics has to be token based, as a matter of principle. The paper investigates the Buridanian considerations leading to these conclusions, and considers some obviously emerging objections to the Buridanian approach.

1. Definition of validity

If we open any modern logic textbook looking for a definition of valid inference, we find something like the following: an inference is valid if and only if it is impossible for the premises to be true while the conclusion is false. At any rate, this formula and its equivalents are usually provided as a first, intuitive introduction of the idea, which then is to be captured in more exact terms in a formal semantic or syntactic system.

It may seem quite surprising, therefore, that in his *Treatise on Consequences*, when John Buridan briefly considers an equivalent formulation of the definition of a valid inference, provided in terms of the definition of the antecedent of a valid consequence, he rejects it. As he writes:

> [The terms] 'antecedent' and 'consequent' are predicated correlatively; therefore, they need to be described in terms of each other. Many people say that of two propositions that one is the antecedent with respect to the other which cannot be true while the other is not true, and that one is the consequent with respect to the other which cannot be not true while the other is true, so that every proposition is antecedent with respect to any other proposition which cannot be true without that other being true. But this description is defective or incomplete, for the following is a valid consequence: 'every man is running; therefore, some man is running';[2] still, it is possible

1 Some of the material contained in this paper will form part of the author's monograph *The Logic and Metaphysics of John Buridan*, forthcoming in the *Past Masters* series of Oxford University Press.

2 Obviously, this consequence is only valid on the medieval analysis of universal affirmatives, attributing 'existential import' to them. For a more detailed discussion of the issue see *Klima* (*2001*, pp. 197–226). Modern readers who still do not like this analysis may substitute their favorite example of a valid consequence here; Buridan's point remains the same.

History and Philosophy of Logic ISSN 0144-5340 print/ISSN 1464-5149 online
http://www.tandf.co.uk/journals
DOI: 10.1080/01445340310001610944

> for the first proposition to be true and for the second not to be true, indeed, for the second not to be.[3]

This perceived defect of the description is clearly a consequence of the common medieval conception of propositions as being singular, contingent, temporary occurrences, whether in writing, in speech, or in the mind.[4] So, apparently, all that is needed to take care of this worry is to supplement the description with a clause stating the requirement of the actual formation of the propositions it concerns. As Buridan continues:

> Therefore, some people say that this description needs to be supplemented as follows: that proposition is the antecedent with respect to another proposition which cannot be true while the other is not true, when they are formed together.[5]

But even this will not do. As Buridan's next argument points out, one cannot rely in this description on the notion of truth at all:

> I still claim that this description is not valid, for the following is not a valid consequence: 'no proposition is negative; therefore, no donkey is running', but on the basis of the given description one should accept it as valid; therefore, etc. I prove the first premise. The opposite of the consequent does not entail the opposite of the antecedent; for this is not valid: 'some donkey is running; therefore, some proposition is negative'. The second premise is clear. For the first proposition, which is designated to be the antecedent, cannot be true; therefore, it cannot be true while the other is not true.[6]

Obviously, the *contrapositive* of the original consequence, namely, 'some donkey is running; therefore, some proposition is negative' is not valid, for in a possible situation a donkey may be running, while God has eliminated all negative propositions. Yet, according to the proposed definition, the original consequence should be deemed valid. For the antecedent 'no proposition is negative' cannot possibly be true, since whenever it is formed, its very existence, being itself a negative proposition, falsifies it. But then, if it cannot be true at all, it certainly

3 Antecedens autem et consequens relatiue dicuntur ad inuicem; ideo per inuicem describi debent. Dicunt ergo multi quod propositionum duarum illa est antecedens ad aliam quam impossibile est esse ueram illa alia non existente uera et illa est consequens ad reliquam quam impossibile est non esse ueram reliqua existente uera, ita quod omnis propositio ad omnem aliam propositionem est antecedens quam impossibile est esse ueram illa alia non existente uera. Sed haec descriptio deficit uel est incompleta, quia hic est bona consequentia omnis homo currit; ergo aliquis homo currit, et tamen possibile est primam esse ueram secunda non existente uera, immo secunda non existente. *Buridan 1976* (henceforth TC) (pp. 21–22)

4 In the context of Buridan's nominalist philosophy, we can safely disregard the issue of *real propositions* espoused by Burleigh and other realists. Cf. *Nuchelmans 1980* (pp. 9–13).

5 Et ideo aliqui dicunt dictam descriptionem debere suppleri sic: illa propositio est antecedens ad aliam propositionem quam impossibile est esse ueram illa alia non existente uera istis simul formatis. TC (pp. 21–22).

6 Sed adhuc dico quod haec descriptio non est bona, quia hic non est bona consequentia nulla propositio est negatiua; ergo nullus asinus currit, et tamen secundum dictam descriptionem oporteret eam concedere esse bonam; ergo etc. Primam praemissam probo. Quia ex opposito consequentis non sequitur oppositum antecedentis; non enim sequitur 'quidam asinus currit; ergo quaedam propositio est negatiua'. Secunda autem praemissa manifesta est. Quia primam, scilicet quae designatur esse antecedens, impossibile est esse ueram; ergo impossibile est ipsam esse ueram alia non existente uera. TC (pp. 21–22)

cannot be true when its consequent is false; therefore, according to the proposed definition its consequent should validly follow from it. However we have just seen that it does not. Indeed, it does not follow even in accordance with the proposed definition. For the denial of the consequent of the original consequence, namely, 'some donkey is running' may be true, while the denial of the antecedent, namely, 'some proposition is negative' is not true in the possible situation in which a donkey is running and God has eliminated all negative propositions. So, since the denial of its consequent does not follow from the denial of its antecedent, the original consequence is not valid, even according to the proposed definition.

Therefore, since the original consequence is clearly not valid, even according to the proposed definition of a valid consequence, whereas according to the same proposed definition it should be valid, that definition leads to inconsistency, and thus it cannot be correct. So, is there something fundamentally wrong with those modern logic text books I referred to in the beginning?

2. Assumptions of Buridan's argument

Let us take a closer look at Buridan's argument. Apparently, its conclusion is that according to the proposed definition of validity the consequence it considers is both valid and not valid. Hence, that definition cannot be correct. Since the conclusion of Buridan's argument does not rely on the proposed definition alone, but on a number of further assumptions, the rejection of the definition is only justified to the extent that those assumptions are justified. So let us see these assumptions. The argument may be spelled out in a somewhat more transparent form as follows:

(1) A consequence is valid if and only if its antecedent cannot be true while its consequent is false, when they both exist [assumed definition]
(2) Any proposition whose existence falsifies it cannot be true [self-evident]
(3) The existence of the proposition 'no proposition is negative' falsifies it [self-evident]
(4) The proposition 'no proposition is negative' cannot be true[7] [2, 3]
(5) The antecedent of 'no proposition is negative; therefore, no donkey is running' cannot be true [4]
(6) The antecedent of 'no proposition is negative; therefore, no donkey is running' cannot be true while its consequent is false, when they both exist [5, $\sim Mp \rightarrow \sim M(p \,\&\, q)$]
(7) The consequence 'no proposition is negative; therefore, no donkey is running' is valid [1, 6]
(8) The proposition 'some donkey is running' can be true while the proposition 'some proposition is negative' is false, when they both exist [self-evident]
(9) The consequence 'some donkey is running; therefore, some proposition is negative' is not valid [8,1]

7 To be sure, given Buridan's token-based conception of propositions, the phrase 'the proposition "p"' does not always manage to single out a single proposition-token (although in some contexts, when the uniqueness of reference is secured by the context itself, it can). For brevity and naturalness of expression, when no confusion arises, we can still use this phrase as an abbreviation of the expression 'any proposition of the form "p"'.

(10) The consequence 'some donkey is running; therefore, some proposition is negative' is the *contrapositive* of the consequence 'no proposition is negative; therefore, no donkey is running' [self-evident]
(11) Whenever the *contrapositive* of a consequence is not valid, then the consequence is not valid [self-evident]
(12) The consequence 'no proposition is negative; therefore, no donkey is running' is not valid [9, 10, 11]
(13) The consequence 'no proposition is negative; therefore, no donkey is running' is valid and it is not valid [12, 7]

So, having arrived at a neat contradiction, we may start checking the auxiliary premises of Buridan's *reductio*.

In the succinct statement of his argument, Buridan only provides the equivalent of what is premise 3 here. In the more extensive parallel discussion in his *Sophismata*, however, Buridan argues for the impossibility of 'No proposition is negative' thus:

> ... the other [proposition] is impossible, namely, 'No proposition is negative', for in no case can it be true. For whenever it is not, then it is neither true nor false, and whenever it is, then some proposition is negative, namely, itself; therefore, it is false to say that no proposition is negative.[8]

That is to say, if a proposition is true or false, it has to exist. But if the proposition 'No proposition is negative' exists, then at least one negative proposition exists, namely, itself, and so it is not true that no proposition is negative. This is precisely what is meant by the claim that the existence of this proposition falsifies it, namely, that its existence implies its contradictory.

If one were to object that this example is a contrived, unique case, the occurrence of which by no means justifies the universal claim made in premise 2, then I can immediately provide an infinity of such self-falsifying propositions. Consider the proposition-scheme 'Every proposition contains *n* words'. Since substituting any numeral in place of *n* in this scheme yields a proposition of five words, we get a false proposition falsified by its own existence if we substitute any numeral in place of *n* other than 5.

Obviously, as I mentioned earlier, Buridan's worries concerning the truth-values of a proposition connected to the existence of a proposition presuppose his conception that propositions are contingent, temporary occurrences. However, do we have a similar situation if we take propositions to be the abstract, atemporal entities expressed by temporally occurring sentence-tokens, as many modern philosophers would take them to be?

If propositions are atemporal, they exist timelessly, that is, there is no time at which they do not exist. So, whenever a timeless proposition is expressed by a temporally occurring sentence-token, then the proposition expressed by that sentence-token exists. Therefore, whenever I form a token of the sentence 'No proposition is negative' the proposition that no proposition is negative expressed by this sentence-token exists. But its existence entails that some proposition is negative, so the proposition cannot be true. Therefore, the atemporal proposition is just as self-falsifying as Buridan's occurrent temporary proposition or the sentence-token is.

8 *Buridan 2001* (henceforth SD) (p. 953).

Clearly, referring to atemporal propositions in the definition of a valid consequence only addresses Buridan's first problem with this definition, namely, the possible non-existence of the proposition that figures as the consequent of an intuitively valid consequence. But the problem of consequences with self-falsifying antecedents is the same for both atemporal and temporal conceptions of propositions.

However, one may also object here to Buridan's procedure of trying to falsify what appears to be an intuitively clear and acceptable definition of validity by challenging the alleged self-evident character of line 11 in the reconstruction.[9] For given Buridan's token-based conception of propositions, that intuitively clear definition immediately invalidates the rule of contraposition, since the existence of the propositions of a consequence is independent from the existence of their negations occurring in the contrapositive. So it is quite possible to have a valid consequence without having the corresponding valid contrapositive, and so the invalidity of the contrapositive of a consequence is not sufficient for the invalidity of the original consequence. But this is precisely what Buridan's argument assumes as self-evident in line 11 of the reconstruction. Therefore, Buridan's argument fails, because assuming the proposed definition and the conception of propositions as contingently existing sentence-tokens, he can no longer assume contraposition (and so the corresponding line 11 in the reconstruction) as self-evident.

To this objection I reply in the first place that although under these assumptions contraposition in general fails, nevertheless, assuming bivalence, and the existence of all four propositions involved in a contraposition, the invalidity of the contrapositive of a consequence is sufficient for the invalidity of the consequence, even according to the proposed definition of validity. In general, if in a possible situation

(a) 'p' exists, and 'q' exists, and '$\sim$p' exists, and '$\sim$q' exists, and '$\sim$q' is true, and '$\sim$p' is false,

then, assuming bivalence, in that possible situation

(b) 'p' exists, and 'q' exists, and 'p' is true, and 'q' is false.

But (a) describes precisely the situation that invalidates '$\sim$q; therefore, $\sim$p' and (b) describes the situation that invalidates 'p; therefore, q', according to the proposed definition. Therefore, line 11 still holds under these assumptions.

However there still is a serious problem with Buridan's particular move in the argument, namely, the application of line 11 to line 9 to get to the conclusion expressed by line 12. For consider again line 8:

(8) The proposition 'some donkey is running' can be true while the proposition 'some proposition is negative' is false, when they both exist [self-evident]

This indeed invalidates the consequence 'some donkey is running; therefore, some proposition is negative' according to the proposed definition, as stated by line 9. But, in accordance with (a) and (b) above, from this we could move to the invalidation of the original 'no proposition is negative; therefore no donkey is running', only if we could assume in the situation described by line 8 the existence of its embedded

9 I owe the gist of this objection to David Kaplan.

propositions, in particular, its antecedent, 'no proposition is negative'. But that is not possible, for if 'no proposition is negative' exists, then 'some proposition is negative' (the consequent of the contrapositive of the original conditional) cannot be false, as described by line 8.

So in this particular case, the situation invalidating the contrapositive (described by line 8) cannot contain the antecedent of the original consequence, and so it cannot invalidate the original consequence *according to the proposed definition*.

However at this point Buridan may ask: why should the validity of the moves in an argument meant to disprove a proposed definition of validity be judged by the standard of the proposed (possibly flawed) definition? After all, we may take it to be a primitive fact about our intuitions of validity that whenever the contrapositive of a consequence is invalid, then the consequence is invalid too, and if a proposed definition of validity does not conform to this primitive intuition, then so much the worse for that proposed definition.[10] Note also that according to Buridan's revised definition (not in terms of truth, but in terms of things being the way they are signified to be) the move in question is valid, because according to that definition the possible situation described by line 8 need not contain the antecedent of the original consequence to invalidate that consequence. So, if the issue boils down to having to choose between whether we should keep the proposed definition and discard contraposition or we should emend the definition so that it accommodates our strong (''pre-definitional'') intuition about the validity of contraposition, then Buridan is at least not unreasonable in choosing the latter alternative.

Thus, if we keep our intuition about contraposition (also validated later by Buridan's emended definition), consequences with self-falsifying antecedents provide a strong case for Buridan's claim that the definition of validity cannot consistently be provided in terms of truth.

3. Tarskian semantics and the 'Reciprocal Liar'

So, given all the trouble these consequences may cause, how come modern accounts of validity do not worry about them at all? The simple answer to this question seems to be that on the modern formal accounts inspired by Tarski's approach to semantics, such consequences, or rather the self-falsifying propositions occurring in them, are inexpressible. Tarski's approach gets rid of all sorts of *insolubilia* by simply banning self-reference from the object language and reserves all semantic discourse for its meta-language. In this way, the definition of a valid inference for the object language can safely be provided in terms of the notion of truth defined in the meta-language, because paradoxes of self-referential expressions, there being no such expressions in the object language, simply cannot emerge. But even if Tarski's approach works for artificial languages of artificially impoverished expressive powers, natural languages would simply not fit into its narrow mold. Natural languages obviously abound with self-referential expressions, indeed, all sorts of indirect self-referential means that give rise to inferences which in principle ought to be able to be captured even by ordinary quantification theory. For example, the inference 'I believe whatever you say; but you say Tarski is right; therefore, I believe Tarski is right' seems to be a pretty straightforward case of universal instantiation in

10 I owe this point to Elizabeth Karger.

which the first premise quantifies over your *dicta*, whatever those are. But if we *can* make such references, as we obviously do all the time, then in natural language semantics we obviously have to be able to handle situations like the following Buridan-inspired case of the "Reciprocal Liar".

Consider the following situation.

Plato says: 'Socrates says something true'.
Socrates says: 'Plato says something false'.
Robert says: 'Plato says something false'.

And they do not say anything else, while both Socrates and Robert think that Plato said something false, namely, that God does not exist.[11]

From these propositions and the situation as described, we can apparently derive the paradoxical conclusion that Socrates' proposition is both true and false by means of the following two arguments.

Argument 1—to show that Socrates' and Plato's propositions are false and Robert's is true. In this case, if Socrates' proposition is true, then Plato's proposition is false. But if Plato's proposition is false, that means that Socrates does not say something true, that is, Socrates' proposition, the only proposition he utters, is false. Therefore, if Socrates' proposition is true, then it is false; so, it cannot be true. However, if Robert's proposition is true, then it does not follow that it is false, indeed, its truth merely entails that Socrates' proposition is false, as has already been established.

Argument 2—to show that Socrates' proposition is both true and false. To be sure, one may believe that this assignment of truth values does not avoid paradox. For if Socrates' proposition that Plato says something false is false, then it is not the case what it signifies to be the case, namely, it is not the case that Plato says something false. But if it is not the case that Plato says something false, and he says something and says nothing else, then, observing the principle of bivalence, he says something true when he says that Socrates' proposition is true; whence Socrates' proposition is true. Therefore, from first to last, if Socrates' proposition is false, then it is true, and, from the previous reasoning, if it is true then it is false; so we have a contradiction.

Buridan's Solution of Argument 2. However, Buridan's theory of self-referential propositions can come to the rescue here. According to Buridan, a true proposition both *formally signifies* that it is the case what it asserts to be the case and *virtually implies* its own truth. So, if a proposition 'p', named A, is true, then it formally signifies that it is the case that p and virtually implies that A is true. In general, the claim that the proposition 'p', named A, is true is equivalent to the claim that it is the case that p and A is true. Therefore, if 'p' is false, then it is false either because it is not the case that p or because A is not true. Now in the

11 See SD (9.8, 8th sophism, pp. 971–974). I have modified Buridan's sophism, since his treatment of it relies on a rather dubious 'parity of reasoning'-style argument attempting to establish the equivalence of 'Plato says something false' and 'Socrates says something false' uttered by Socrates and Plato respectively. I believe the modified version serves to illustrate the points I want it to illustrate without having to rely on this type of reasoning.

case of Socrates' proposition, if we call it A, this means that if it is false, then it is either false because it is not the case that Plato says something false, or because A is not true. But on account of the first argument [Argument #1] we have seen that A is not true, because its truth entails its falsity. So Socrates' proposition is false not because it is not the case what it formally signifies to be the case, but because it is not the case what it virtually implies to be the case, namely, that it is true. Therefore, the falsity of Socrates' proposition does not entail the truth of Plato's proposition; it merely entails its own falsity. So, Socrates' proposition is false, and so is Plato's, whence Robert's is true, without any contradiction.

4. Logic without truth, and the logic of 'true'

As J. L. Austin allegedly once remarked concerning Tarski's biconditional ('Snow is white' is true if and only if snow is white): so far, so good.[12] Buridan's approach successfully avoids paradox by blocking the apparent 'reverse implication' of Liar-sentences, namely, the implication from their falsity to their truth; whence they can simply be regarded as false. Surprisingly, however, Buridan achieves this by disagreeing with Austin: despite its apparent plausibility, the Tarski biconditional is not good for Buridan at all. It is not good precisely because it would license the paradoxical 'reverse implication', and this is because, according to Buridan, something important is missing on the right-hand-side of this biconditional. (For if 'A is false' is true if and only if A is false, where A is the name of the sentence in quotes within this sentence, then—moving from the right-hand-side of the biconditional by means of what I called the 'reverse implication'—if A is false, then A is true.)

In his *Treatise on Consequences*, immediately after stating what at first may appear to be a list of truth-conditions for various types of categorical sentences in terms of the supposition (i.e. reference) of their terms (e.g. that an affirmative proposition designates that its terms supposit, i.e. stand, for the same thing or things, while the corresponding negative designates the opposite, etc.), Buridan makes the following remark:

> Although these points are conceded, nevertheless, it does not follow that every assertoric, present tense, affirmative categorical proposition the terms of which supposit for the same is true. For in the case of a proposition asserting itself to be false, it is possible that its terms supposit for the same and yet it is false, for example, if someone says: 'The proposition I am uttering is false'. And the reason is that although this proposition on account of its form designates that its terms supposit for the same, and that is indeed the case,[13] nevertheless, on account of the signification of its predicate it designates that they do not supposit for the same. For when we claim a[n affirmative] proposition to be false, we designate that its terms do not [supposit for the same]. Therefore, a proposition like this designates that [the things supposited for by its terms] are the same and that they are not the same,

12 I owe this anecdotal remark to Peter King. As in almost all cases of such anecdotal remarks, the Italian saying applies here too: *se non è vero è ben trovato*.

13 The terms 'the proposition I am uttering' and 'false' do indeed stand for the same thing, namely, the proposition I am uttering, given that it is in fact false.

> and so, although things are in the way in which it signifies them to be, yet, they are not in every way in which it signifies them to be, and therefore it is false.[14]

The crucial idea here is that a Liar-sentence must be false not because it is not the case what it designates to be the case on account of its form [*de sua forma*], but because somehow it also designates the opposite. Buridan does not elaborate this point here, and, in fact, the close resemblance of this formulation to what we find in his *Questions on the Posterior Analytics* may indicate that this formulation expresses a view that Buridan held earlier, but revised in his full account of *insolubilia*, in chapter 8 of his *Sophismata*:

> For some people have said, and so it seemed to me elsewhere,[15] that although this proposition does not signify or assert anything according to the signification of its terms other than that every proposition is false, nevertheless, every proposition by its form signifies or asserts itself to be true. Therefore, every proposition asserting itself to be false, either directly or implicitly, is false, for although things are as it signifies, insofar as it signifies itself to be false, nevertheless, things are not as it signifies insofar as it signifies itself to be true. Therefore, it is false and not true, since for its truth it is required not only that things be as it signifies but also that they be in whatever way it signifies [them to be]. But this response does not seem to me to be valid, in the strict sense.[16]

After Buridan disposes of this formulation by means of a brief argument,[17] he states what he finds to be a more satisfactory formulation:

> Therefore, we put this otherwise, in a manner closer to the truth, namely, that every proposition virtually implies [*implicat virtualiter*] another proposition in which the predicate 'true' is affirmed of the subject that supposits for [the original proposition]; and I say 'virtually implies' in the way in which the antecedent implies that which follows from it. Therefore, a proposition is not true if in such an affirmative consequent the subject and the predicate do not supposit for the same [thing or things]. For example, let us posit that Socrates' proposition 'No

14 Et licet haec concedantur, tamen non sequitur quod omnis propositio affirmatiua de inesse et de praesenti sit uera cuius termini supponunt pro eodem, quia in propositione asserente se esse falsam potest esse quod termini supponant pro eodem, et tamen ipsa est falsa; uerbi gratia, si aliquis dicat 'propositio quam ego profero est falsa'. Et causa est quia, quamuis illa propositio de sua forma designet idem esse pro quo termini supponunt et ita sit, tamen cum hoc, propter significationem praedicati, designat quod non sit idem. Quamcumque enim propositionem dicimus esse falsam designamus non esse idem pro quo etc. Ideo talis propositio designat esse idem et non esse idem, et ideo, licet qualiter significat ita sit, tamen non qualitercumque significat ita est, et ideo est falsa. TC (c. 5, pp. 25–26)

15 J. Buridan, *Quaestiones in primum librum Analyticorum Posteriorum*, q. 10 (unpublished edition by H. Hubien). Buridan here apparently took over Thomas Bradwardine's solution. Cf. *Read 2002* (pp. 189–218). For more detailed accounts of Buridan's parallel passages see *Pironet 1993* (pp. 293–300); *Buridan 1994* (Introduction §3).

16 SD (pp. 967–968).

17 The point of the argument is that the proposition 'the proposition "p" signifies itself to be true' cannot be true. For according to Buridan's theory of sentential nominalisations, the phrase 'itself to be true' can be taken to stand either materially, for the proposition 'the proposition "p" is true' (but if 'p' is a proposition about things that are not propositions, then it certainly does not signify this proposition, despite what the original proposition claims in this sense), or personally, for whatever the terms of the phrase 'that p is true' are jointly true of, which in the case of an impossible proposition cannot be anything, and thus cannot be signified by the proposition 'p'. Cf. SD (p. 969, n. 183). So, Buridan actually had a very good reason to reject Bradwardine's (and his own earlier) solution: the solution provided in terms of signification cannot be expressed in a true sentence.

> proposition is true' is named by the proper name C; then this is valid: 'No proposition is true; therefore, C is true'; therefore, things are not as the consequent virtually implied by the original proposition signifies, whence that original proposition is not true. For it is not sufficient for the truth of the proposition that things be as it signifies by its formal signification, but it is also required that things be as its virtually implied consequent signifies. And this is why it is said that when a proposition has or can have reference to itself [*habet vel habere potest reflexionem super se*], it does not suffice for the truth of an affirmative [proposition] that its terms supposit for the same [thing or things], as I have said elsewhere,[18] but it is also required that even in this consequent the terms supposit for the same [thing or things], and then it is necessary, if this holds, that the proposition be true.[19]

So, according to this formulation, the truth of a proposition requires not only that it be the case what the proposition formally signifies to be the case, but also that it be the case what it virtually implies to be the case, namely, that the term 'true' be verified of it. Accordingly, Buridan's doctrine seems to be committed to the equivalence I stated above in the reconstruction of the Buridanian solution of the ''Reciprocal Liar'':

> (TB) The proposition 'p', named A, is true if and only if it is the case that p and A is true.

However if we look at this formulation, especially with an eye to the apparently very similar Tarskian biconditional ((T) 'p' is true if and only if [it is the case that] p), then a number of questions arise.

For what is this equivalence supposed to tell us? It is certainly not a 'definition of truth' in a Tarskian, or for that matter, Aristotelian sense, for as a definition it would obviously be circular in *any* sense. Nor does it seem to provide some other sort of explanation of the meaning of the term, or some other theoretical insight into its role in logic or natural language reasoning, as for example modern deflationary theories do. So what is it that it *does* claim? Why is it important? And why is it not a mere *ad hoc* device for blocking the 'reverse implication' of Liar-type sentences, for simply dodging paradox wherever it emerges?

Buridan gives us few clues to answer questions of this kind, nevertheless, here I attempt to provide some answers that I find at least acceptable and in accord with Buridan's relevant remarks. Clearly, (TB) is not a definition of truth. Indeed, Buridan does not appear to be interested in giving such a definition, and, at any rate, his logic certainly does not need one. After all, we have just seen that he finds the central logical notion of validity undefinable in terms of the notion of truth. But then, a definition of truth cannot and need not play the theoretical role in his logic it does in logics with a truth-based notion of validity. On the other hand, what (TB) does say in the first conjunct of its right-hand-side is that for the correct assertion that 'p' is true it is required that 'p' satisfy a certain 'correspondence-condition', namely that things are in the way it signifies them to be. But, as Buridan points out, it is precisely this sort of condition that we need to take into account in formulating the correct definition of validity, and not the truth of the proposition, for a proposition may adequately correspond to a possible situation in which it does not exist, and hence in which it is

18 SD (p. 857).
19 SD (p. 969).

not true. In fact, once we become accustomed to the idea of talking about propositions in the sense of contingently existing proposition-tokens, this is a very natural idea. For example, the proposition 'Dinosaurs were roaming the earth before man appeared' certainly corresponds to a situation in which it did not exist, and hence in which it was not true. But then, with this conception in mind, all we need for validity is this idea of correspondence on the part of our actually formed propositions, which may nevertheless correspond to possible situations in which they do not or even cannot exist. So Buridan's logic, construed strictly as the theory of validity of consequences, could in principle dispense with the notion of truth altogether.

On the other hand, there still is a need for a logical account of the behavior of the term 'true', especially given the fact that it is by means of this term that we normally express our evaluation of propositions for their correspondence with reality. For if we say that a proposition is true, then we want to express that things are as it signifies them to be. Yet, as we could see, just because things are as a proposition signifies them to be, it does not necessarily follow that the proposition is true, for the situation described by the proposition may be such that the proposition does not or even cannot exist in it. Indeed, even if the proposition in question exists in the given situation, but, perhaps, along with a description of the circumstances of the situation, it entails that it is not true, it cannot be true, even if things are in that situation as it signifies them to be. This is precisely the case with Liar-type propositions. Therefore, it should be clear that a proposition cannot be called true on the basis of the satisfaction of its correspondence-conditions alone (namely, things being as it signifies them to be), even if its correspondence is certainly *required* for its truth. What is, then, *sufficient* for the truth of a proposition?

So far, it is clear that the correspondence of the proposition with a given situation and the proposition's existence in that situation are required for the proposition's being true in that situation. It is also clear that if the proposition is to be true in that situation, then it cannot entail that it is not true in that situation.

It is at this point that Buridan's idea of the 'virtual implication' of the truth of a proposition becomes relevant. The point of the qualification 'virtual' in this phrase seems to be that for this implication to hold the consequent need not actually be formed, although, of course, the antecedent has to exist (for otherwise it could not be said to be true).[20] For instance, in the situation in which Socrates only says 'Socrates says something false' and nobody says, thinks or writes anything else, Socrates' proposition virtually implies that Socrates' proposition is true, even if nobody forms the proposition 'Socrates' proposition is true', and so this proposition is not actually implied by Socrates' proposition. However, it is also clear that in the situation in question it is *not* the case what this virtually implied proposition *would* signify to be the case, if it *were* formed. So in this situation things are not as the virtually implied proposition signifies them to be. But then it is clear that Socrates' proposition cannot be true. Yet, curiously, precisely for this reason, things are the way Socrates' proposition signifies them to be, for it says that Socrates' proposition is false.

This situation, then, provides us with an interesting lesson concerning self-referential propositions that falsify themselves. For in their case, and in their case alone, a strange thing happens: even if they adequately correspond to the situation they formally signify to be the case, they cannot be true, because of the failure of the correspondence of the proposition they virtually imply, stating their truth. Then this

20 Cf. SD (9.8, p. 970).

has to be the additional condition required for the truth of any proposition which these propositions, and only these propositions, fail to satisfy, whereas all other propositions trivially satisfy, when their correspondence condition is satisfied. Therefore, in general, we can correctly call a proposition 'true' only if it meets both conditions, namely both its own correspondence-condition and the correspondence-condition of its virtually implied proposition. But this is exactly what (TB) stipulates. In short, 'p and ''p'' exists' entails ' ''p'' is true'.[21] But ' ''p'' is true' entails ' ''p'' exists', as well as 'the correspondence-conditions of ''p'' are satisfied', and 'p' itself. But, the latter together entail ' ''p'' is true'. Therefore, ' ''p'' is true' is equivalent to 'the correspondence-conditions of ''p'' are satisfied and ''p'' is true' (which amounts to (TB) above). This is not a theory of truth or an explanation of what 'true' means, but a trivial equivalence falling out from equally trivial considerations concerning the assertibility conditions of the predicate 'true'. But it does not hurt that it helps avoid Liar-type paradoxes.

So (TB), which is, again, neither a definition, nor a ''theory of truth'', can be regarded simply as a general claim specifying the conditions for the correct assertion of the predicate 'true' of a proposition. The conditions it specifies are trivial, and indeed, in ''normal'' cases they necessarily obtain or fail to obtain together. Yet, since in the case of (whether direct or indirect) Liar propositions the second condition (correspondence of the virtually implied proposition) fails to obtain precisely because the first condition (the correspondence of the proposition itself) obtains, such propositions can never be correctly said to be 'true', whence, observing bivalence, they always have to be called 'false', without the risk of any paradox.[22]

5. Two conclusions of Buridan's solution

I take it that the foregoing discussion establishes two fundamentally important conclusions concerning natural language semantics.

(1) Natural languages are essentially semantically closed, yet they do not have to be on that account inconsistent.
(2) Natural language semantics has to be token based, as a matter of principle.

21 This is the principle G. E. Hughes called Buridan's 'entailment principle', which he ingeniously defended against an apparently obvious objection in the Introduction of his *John Buridan on Self-reference* (*Hughes 1982*, pp. 23–27). (Hughes' defense of this principle, by the way, actually relies on considerations similar to those concerning the situation of the 'Reciprocal Liar', pointing out that two tokens of the same proposition-type may have opposite truth-values. So, I think Hughes is absolutely correct in remarking that Buridan had everything at his disposal to reply to the objection along the lines Hughes does.) As in a stimulating discussion of an earlier draft of this paper at Boise State University Tony Roark has insightfully pointed out, a similar objection raised by Andrew Cortens, attempting to refer to the proposition in the antecedent of this principle, can be handled by relying on Buridan's conception according to which propositional components of complex propositions are not themselves propositions. For more on Buridan's conception, see *Klima 2004*.

22 A similar, yet slightly different, account is provided by *Hughes 1982* (pp. 22–23). Indeed, perhaps surprisingly, I also agree here with Stephen Read's (*2002*, p. 202) conclusion: 'Buridan ends up with no theory of truth at all'. Yet, *pace* Read, I do not find this to be a fatal flaw of Buridan's approach. For on my understanding (TB) does not even attempt to be a theory of truth, and Buridan does not even need such a theory. He has a theory of correspondence, which is all he needs for his logic, and he provides the trivial assertibility conditions for the predicate 'true' in his token-based semantics, whereby he can show why, despite possible appearances to the contrary, what Tarski would call the 'semantic closure' of his theory does not entail paradoxical results.

Pro 1. Since natural languages primarily evolved for communication, and not for solitary reasoning, in natural languages we routinely need to refer to and judge other people's judgments, which in turn, even through several intermediaries, may concern our own. In short, natural languages are what Tarski called *semantically closed* languages. So, even if we were to ban direct self-reference by some sort of weird legislation, indirect self-reference would be inevitable in natural languages. Yet, as Buridan's solution shows, this situation need not entail that semantic paradoxes are also inevitable in natural languages. In general, self-referential claims that entail whether directly or indirectly their own falsity are false. Still, Buridan's plausible requirement for their truth, namely, that besides the obtaining of what they formally signify, the virtual implication of their own truth should also hold, effectively blocks the further, paradoxical inference of their truth from their falsity. So they are simply false, and their falsity does not entail their truth—there is no paradox.

Pro 2. In the case of the Reciprocal Liar, Socrates and Robert form distinct tokens of the same proposition-type, not only in speech or writing, but even in intention. But precisely on account of their being distinct tokens of the same type, these propositions have opposite semantic values, that is, radically different semantic properties. However, from the occurrence of this phenomenon we have to conclude that natural language semantics primarily has to concern token-expressions and not types, for different tokens of the same type may have different semantic properties to be accounted for in logical theory. Therefore, even apart from the metaphysical and theological reasons against the dubious ontology of abstract propositions provided by Buridan in his refutation of the theory of *complexe significabilia*, this sophism seems to present us with a purely logical reason to prefer a primarily token-based semantics to a semantics that only concerns itself with expression-types.[23]

6. Consequences of a token-based semantics

But what does this really mean? Can a semantic theory *not* concern itself with expression-types rather than tokens? After all, in such a theory we are supposed to formulate universal logical laws, and so these laws, as such, are supposed to concern types rather than tokens, aren't they?

Of course, his staunch nominalism notwithstanding, even Buridan would agree that logical laws must be universal. But he would insist that a universal law is not supposed to be universal because it concerns universal entities, but because it equally and universally applies to several singulars. So, if we understand talking about types not as discourse about some universal, abstract entities, but as a form of speech that must be interpreted as equally concerning several singulars together on account of their similarity, then Buridan would certainly not object—after all, as he insists, *nomina sunt ad placitum*.

Indeed, Buridan's well-known liberalism concerning usage is at work even in his further discussion of the proposed definition of the antecedent of a valid consequence.

23 Cf. *Gaifman 2000* (pp. 79–121). It is quite revealing that Gaifman's main motivation for a token-based semantics is the same sort of semantic paradox that I dubbed the 'Reciprocal Liar' in Buridan.

Having rejected the definition provided in terms of truth, he first proposes his own, revised definition, not relying on the notion of truth. As he writes:

> Therefore, others define [antecedent] differently, [by saying that] that proposition is antecedent to another which is related to it in such a way that it is impossible for things to be in whatever way the first signifies them to be without their being in whatever way the other signifies them to be, when these propositions are formed together.[24]

However Buridan is not entirely satisfied even with this formulation. As he goes on:

> However, this description is not true by virtue of the proper force of the expression [*de virtute sermonis*], because it presupposes that any proposition is true because things are in whatever way it signifies them to be, which we denied earlier. But then we also said that we were going to use this mode of expression in the sense given earlier; therefore, in that sense we accept this description.[25]

Buridan here refers to his earlier discussion of the truth-conditions of various types of propositions. There he established that one cannot provide a single formula for all types of propositions, but different formulations should be provided for different types. Still, in the end he did accept a single formula with the proviso that it should be understood as a sort of shorthand for those different formulations for all different types.[26] Indeed, with a similar proviso in mind, he is even willing to accept the proposed definition of the antecedent of a valid consequence in terms of truth. As he continues:

> Indeed, we are often going to use the mode of expression in accordance with the first description [of antecedent] manifestly disproved above, for it has counterexamples only in the cases of a few consequences. But whatever mode of expression we shall use, we shall always intend the sense given above.[27]

24 Ideo alii aliter diffiniunt, dicentes quod illa propositio est antecedens ad aliam quae sic se habet ad illam quod impossibile est qualitercumque ipsa significat sic esse quin qualitercumque illa alia significat sic sit ipsis simul propositis. TC (c. 3, p. 22)

We should also keep in mind the crucial distinction concerning the correct interpretation of this rule Buridan introduced in his *Sophismata*: '. . . a consequence is valid if it is impossible for things to be as the antecedent signifies without their being as the consequent signifies. And this rule can be understood in two ways: first, that it is one proposition about impossibility in the composite sense, in the way that this is commonly used, and its sense then is that this is impossible: 'When it is formed, things are as the antecedent signifies and not as the consequent signifies'. And taken in this way the rule is not valid [. . .]. Taken in the other way, the rule is understood as a proposition about impossibility in the divided sense, so that its sense is: 'a consequence is valid if in whatever way the antecedent signifies [things to be], it is impossible for things to be in that way without their being in the way the consequent signifies [them to be]' SD (9.8. *Second sophism*, pp. 957–958).

25 Tamen adhuc illa descriptio non est uera de uirtute sermonis, quia supponit quod omnis propositio uera ex eo sit uera quia qualitercumque significat ita est, quod prius negatum est. Tamen dictum fuit quod hoc modo loquendi uteremur ad sensum prius datum; ideo sic illam descriptionem concedemus. TC (p. 21)

26 See TC (c. 1, pp. 18–19). Cf. SD (9. 2, 14th conclusion, pp. 858–859).

27 Immo etiam saepe utemur modo loquendi secundum primam descriptionem prius manifeste improbatam, quia ipsa in paucis consequentiis habet instantiam. Tamen quocumque modo loquendi utemur nos intendemus sensum praetactum. TC (c. 3, p. 22).

7. Conclusion

All in all, we can summarise Buridan's approach to consequences in the following way. Consequences, just as any other types of propositions,[28] are singular, contingent occurrences, whether in speech, in writing, or in the mind. Indeed, primarily in the mind, since those in writing or speech are propositions only insofar as they are subordinated to those in the mind. This has to be the case not only on the grounds of general metaphysical considerations concerning the dubious ontology of alleged abstract propositions, but as a matter of logical principle. For given the direct as well as indirect self-referential capabilities of natural languages, it can be shown that different token expressions of the same type can have different logical properties on account of being different tokens, one with, and another without, indirect reference to itself.

However this conception of propositions has two, somewhat surprising consequences in the theory of consequences. In the first place, given the contingency of proposition-tokens, in the definition of validity, which should concern all possible pairs of propositions related as antecedent and consequent, one has to take note of the possibility that one proposition of the pair may exist without the other in a possible situation, so the truth of the antecedent alone may not guarantee the truth of the consequent, since the existence of the antecedent does not guarantee the existence of the consequent. In the second place, even more surprisingly, the definition of validity cannot in principle be based on the definition of truth, for given the semantic closure of natural languages that definition leads to paradox.

Nevertheless, an alternative formulation of the definition of valid consequence avoiding the paradox can be provided. Indeed, that alternative formulation is equivalent to the formulation in terms of truth in all cases that do not involve direct or indirect self-reference. Therefore, even that formulation can safely be relied on, as long as we keep in mind that, as a matter of principle, in natural languages that simple formulation has to be understood as a simplified and improper expression of a significantly more complicated truth.

Acknowledgements

Earlier drafts of this paper were presented at the *Moody Conference in Medieval Philosophy* on 'Truth and Proof in the Middle Ages' at UCLA, on 1 February 2003 and in the *Philosophy Colloquium Series of Boise State University*, Idaho, 3 May 2003. I am grateful to both audiences for many insightful comments, in particular to Andrew Cortens, David Kaplan, Elizabeth Karger and Tony Roark. But I want to thank especially Calvin Normore for the numerous inspiring conversations we had on the subject during my visit at UCLA's Philosophy Department during the academic year 2002–2003. Research for this paper was supported by an ACLS Fellowship and a Fordham Faculty Fellowship for the same period.

References

Buridan, J. 1976. *Tractatus de Consequentiis* (TC), H. Hubien, ed., Philosophes Médiévaux, vol. 16. Louvain: Publications universitaires.

Buridan, J. 1994. *Questiones Elencorum*, R. van der Lecq and H. A. G. Braakhuis, eds, Nijmegen: Ingenium Publishers.

28 Concerning Buridan's treatment of consequences as a special type of hypothetical propositions in which the clauses are asserted, see *Klima 2004*.

Buridan, J. 2001. *Summulae de Dialectica* (SD), tr. intro. and notes G. Klima, New Haven: Yale University Press, 2001.

Gaifman, H. 2000. 'Pointers to Propositions', in Andre Chapuis and Anil Gupta, *Circularity, Definition, and Truth*, Indian Council of Philosophical Research, New Delhi/Atascadero, CA: Indian Council of Philosophical Research/Ridgeview, pp. 79–121.

Hughes, G. E. 1982. *John Buridan on Self-reference*, Cambridge: Cambridge University Press.

Klima, G. 2001. 'Existence and Reference in Medieval Logic', in A. Hieke and E. Morscher, *New Essays in Free Logic*, Dordrecht: Kluwer Academic Publishers, pp. 197–226.

Klima, G. 2004-forthcoming. ''John Buridan and the Force-Content Distinction', in A. Maierù and L. Valent, *Medieval Theories on Assertive and Non-Assertive Language*, Acts of the 14th European Symposium on Medieval Logic and Semantics, Rome: Olschi.

Nuchelmans, G. 1980. *Late-Scholastic and Humanist Theories of the Proposition*, Amsterdam-Oxford-New York: North Holland Publishing Company.

Pironet, F. 1993. 'John Buridan on the Liar Paradox: Study of an Opinion and Chronology of the Texts' in K. Jacobi, *Argumentationstheorie*, Leiden: Brill Academic Publishers, pp. 293–300.

Read, S. L. 2002. 'The Liar Paradox from John Buridan back to Thomas Bradwardine', *Vivarium* **40**, 189–218.