

# *Semantic defectiveness and the liar*

**Bradley Armour-Garb & James  
A. Woodbridge**

## **Philosophical Studies**

An International Journal for Philosophy  
in the Analytic Tradition

ISSN 0031-8116

Philos Stud

DOI 10.1007/s11098-012-9915-6



 Springer

**Your article is protected by copyright and all rights are held exclusively by Springer Science+Business Media B.V.. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your work, please use the accepted author's version for posting to your own website or your institution's repository. You may further deposit the accepted author's version on a funder's repository at a funder's request, provided it is not made publicly available until 12 months after publication.**

# Semantic defectiveness and the liar

Bradley Armour-Garb · James A. Woodbridge

© Springer Science+Business Media B.V. 2012

**Abstract** In this paper, we do two things. First, we provide some support for adopting a version of the *meaningless strategy* with respect to the liar paradox, and, second, we extend that strategy, by providing, albeit tentatively, a solution to that paradox—one that is *semantic*, rather than *logical*.

**Keywords** Truth · Liar paradox · Pretense · Expressibility

## 1 Pretense and meaninglessness

The starting point for our approach to the liar paradox, and, in fact, to cases of semantic pathology generally, will be our pretense account of truth-talk (see Woodbridge 2005 and Armour-Garb and Woodbridge 2010). However, while we favor couching our proposal in the context of the pretense view, our approach to liar sentences and similar phenomena really makes use of certain consequences of that view and does not directly appeal to its details. So, our proposal for the liar paradox should be available to other accounts of truth-talk with the same relevant consequences.

Our approach is a version of the meaningless strategy, according to which liar sentences turn out to lack content in a certain sense. Since a central claim of the pretense account of truth-talk is that, really, there is no property of truth, we do not hold that the content of any sentence is constituted by, or explained in terms of, its truth-conditions. We do, however, maintain that there is an important sense of content that a sentence can have that involves the sentence specifying conditions

---

B. Armour-Garb (✉)  
University at Albany-SUNY, Albany, NY 12222, USA  
e-mail: armgrb@albany.edu

J. A. Woodbridge  
University of Nevada, Las Vegas, Las Vegas, NV 89154-5028, USA  
e-mail: woodbri3@unlv.nevada.edu

that can obtain or not. We call such conditions, *M-conditions*. Truth-conditions are related to M-conditions, in some respects, but the former have only a thin, derivative status, as conditions for the appropriate application of the truth-predicate. On our view, the truth-conditions for a sentence are a *by-product* of its meaning, of which M-conditions are a significant component. This is in line with the meaning-to-truth conditional,

(MTC) If S means that p, then S is true iff p,

no instance of which we reject.

Now, while some sentences specify M-conditions directly, as is the case with

(1) Snow is white,

others specify M-conditions only indirectly. Indeed, one of the consequences of our pretense account of truth-talk is that any specification of M-conditions (that obtain or fail to obtain outside of the pretense) that is accomplished by a 'true'-involving sentence will be accomplished only indirectly, via the operation of the pretense that governs the functioning of the truth-predicate. In particular, it turns out that any M-conditions specified by an instance of truth-talk must be a function of conditions specified by the supposed content-vehicle that is putatively denoted in that instance of truth-talk.

To see this, consider a straightforward instance of truth-talk, such as

(2) 'Snow is white' is true.

When asserted, a 'true'-involving sentence like (2) presents the pretenses it displays as *prescribed*, where being prescribed is determined by:

- (a) the particular principles of generation that govern those pretenses, and
- (b) whether the M-conditions, whose obtaining those principles make prescriptive for the pretenses, actually obtain.

The particular principles of generation governing the operation of the predicate 'is true' determine that the prescriptive M-conditions for the pretenses displayed in (2) are those specified by the content-vehicle that (2) purports to denote.<sup>1</sup> So, by presenting the pretenses it displays as prescribed, (2) specifies, indirectly, precisely the M-conditions that (1) specifies directly.

<sup>1</sup> Although we shall not be arguing for the set of principles of generation that we prefer, for the sake of completeness, we list them here:

- (i) It is to be pretended that expressions like 'is true' and 'is false' function predicatively to describe objects as having or lacking properties (called, not surprisingly, 'truth' and 'falsity').
- (ii) The pretenses displayed in an utterance of  $\ulcorner \langle p \rangle$  is true  $\urcorner$  are prescribed iff p.
- (iii) The pretenses displayed in an utterance of  $\ulcorner \langle p \rangle$  is false  $\urcorner$  are prescribed iff  $\sim p$ .
- (iv) If  $S_1$  and  $S_2$  are sentences that are alike except that (in some transparent context), one has as a subsentence  $\ulcorner p \urcorner$  where the other has  $\ulcorner \langle p \rangle$  is true  $\urcorner$ , then one can directly infer  $S_1$  from  $S_2$  and vice versa.

For more on these principles, see Woodbridge (2005) and Armour-Garb and Woodbridge (2010).

As we will show, this account of truth-talk has an interesting consequence for liar sentences and their kin: They do not specify any M-conditions. In the case of a liar sentence, such as

(L) (L) is not true,

any M-conditions that (L) specified would have to be a function of the M-conditions specified by the content-vehicle that this instance of truth-talk putatively denotes. But in this case that is “another” instance of truth-talk (in fact, it is (L) itself). This means that in order to determine the M-conditions that (L) would specify, we must look to what content-vehicle this “other” instance of truth-talk putatively denotes.

In some cases, this iterated process will “ground out,” as it does for

(3) Sentence (2) is not true.

Any M-conditions specified by (3) indirectly would be a (negating) function of the M-conditions specified by (2). Since (2) is itself an instance of truth-talk, any specifying of M-conditions that it accomplished would also happen only indirectly, as a function of the M-conditions specified by the content-vehicle that it putatively denotes—here, the M-conditions that (1) specifies directly.

Unlike in the (3)–(2)–(1) case, however, with a liar sentence like (L), this multi-step determination process repeats endlessly, with the result that (L) never manages to specify any M-conditions. In a sense, we get instructions that can never be completed.<sup>2</sup> To see this, recall that to determine what M-conditions (L) would specify, the principles of generation for the pretense governing the use of the truth-predicate tell us to look to the content-vehicle that the truth-predicate gets applied to in (L). When we check this content-vehicle, to determine what M-conditions it specifies, we see that it too is an instance of truth-talk—specifically, (L) itself. Accordingly, in order to determine the M-conditions that this content-vehicle specifies, we have to apply the principles of generation for the truth-predicate again, which, again, tell us to look at what M-conditions get specified by the content-vehicle now putatively being denoted in this “new” instance of truth-talk. But because this content-vehicle is, again, just (L) itself, the process repeats and repeats without ever grounding out. So, the steps dictated by the principles of generation for ‘is true’ never finish, with the result that no M-conditions ever get determined as prescriptive for the pretenses displayed in (L).<sup>3</sup>

Since (L) and, indeed, liar sentences and their kin more generally, do not specify any M-conditions as prescriptive for the pretenses they display, there are no

<sup>2</sup> The same goes for wide-scope negation liar sentences, e.g.,  
(L\*) It is not the case that (L\*) is true.

The pretense never determines prescriptive M-conditions for these sentences either.

<sup>3</sup> For slightly different reasons, a similar diagnosis will apply to the following case, suggested by an anonymous referee:

(O) The M-conditions of (O) do not obtain.

The self-referential aspect of this case creates a (slightly different) kind of looping, one that would require (O) to have M-conditions *already*, in order for any M-conditions to be determined for it. But this kind of semantic bootstrapping is impossible, so (O) fails to specify any M-conditions as well. Thanks to the referee, for the helpful suggestion.

M-conditions the obtaining of which one can indirectly commit oneself to, by putting forward a liar sentence. Accordingly, in the “specification of M-conditions” sense of content that we intend here, no content ever manages to attach to a liar sentence like (L). Thus, liar sentences end up not having any content. We therefore conclude that, *in a sense*, they are meaningless. As such, we presently endorse a version of the “meaningless strategy” for dealing with the liar paradox.<sup>4</sup>

## 2 Meaninglessness and understanding

Any meaningless strategy faces an immediate objection, which arises once we recognize that it seems unequivocal that, in some sense, we *understand* liar sentences.<sup>5</sup>

Here is the objection, given that we seem, at least *prima facie*, to understand liar sentences: It is an apparent platitude that to understand a sentence is to know what it means. Moreover, in order to know what a sentence means, it has to be meaningful. Thus, on the assumption that we understand a liar sentence like (L), it seems to follow that (L) is meaningful, contrary to what meaningless strategists have held.

In response, we explain that we do not deny that we can understand a liar sentence like (L), but it is important to note that when we do we only understand (L) *in a sense*. Our claim is that there are (at least) two modes of understanding and that, while we understand (L) in one sense, we do not understand it in another. Call the sense in which we do *not* understand (L), the sense that would require knowing what M-conditions (L) specifies, ‘understanding<sub>1</sub>’. Call the sense in which we do understand (L) ‘understanding<sub>2</sub>’.

To grasp this latter mode, consider an indexical-containing sentence whose referent is unidentified (alternatively, consider Strawson’s (1950) example, ‘This is a fine red one’, where ‘this’ does not have any referent).

So, for example, suppose that you see, written on a whiteboard somewhere, the sentence

(4) I am here now,

though when it was written or by whom are both utter mysteries. Now, clearly a token like (4) is not nonsense. Moreover, although you do not know who the writer is nor when she wrote what she did, there is a sense in which you understand the sentence. But the mode of understanding that you possess is primarily *metalinguistic*: You know what the sentence appears to “say”, for example, that it is of the form: ‘I am here now’, with ‘I’ as the subject, in nominal position, etc. And you know the meaning—the *character*, though not the *content*—of the expressions

<sup>4</sup> This diagnosis, and the treatment we offer below, also applies to truth-teller sentences, such as (K) (K) is true,

which appear to manifest indeterminate semantic pathology.

<sup>5</sup> As an aside, we should note that perhaps the main objection here arises as a consequence of so-called “contingent liars”. Since the points that we will make here clearly apply to contingent liars, we will not address them directly in this paper. For more on contingent liars, see Armour-Garb (2001).

contained therein. Finally, you know how such a sentence could be used to make a meaningful assertion.

Now, our claim is that if you know the form of the sentence, the meanings of the words that are contained therein and how the sentence could be used to make a genuine assertion, then you can be said to “understand<sub>2</sub>” the sentence.<sup>6</sup> But insofar as you do not know the M-conditions for (4), or even whether there are any, then, while you may—indeed, probably do—understand<sub>2</sub> sentence (4), you do not understand<sub>1</sub> that sentence. Of course, this is not to say that a token of (4) could not have M-conditions. But even if it did, as things stand, you do not understand<sub>1</sub> that sentence.

Before moving on, we should note that we are *not* claiming that the sense in which we do not understand (4) is exactly analogous to the sense in which we do not understand (L). Rather, we are providing a rather neutral case in which it seems appropriate to say that one understands (4) in one sense, while denying that she understands (4) in another sense, as a means for explaining the two modes of understanding, understanding<sub>1</sub> and understanding<sub>2</sub>.

## 2.1 Understanding and liar sentences

Having briefly sketched these two modes of understanding, we turn to consider liar sentences. We contend that although we understand<sub>2</sub> liar sentences, we do not understand them in the more substantive sense of understand<sub>1</sub>, since they fail to specify M-conditions and, thus, are meaningless in the way that we have indicated. That said, liar sentences do appear to be meaningful, in a way that underwrites the understanding<sub>2</sub> we have of them.

For example, in the case of (L), we know that the predicate ‘is true’ is being applied to some putative content-vehicle denoted by the subject expression, and we know how such a sentence could be used to make a genuine assertion (e.g., if the sentence name, ‘(L)’, were to name a sentence like (1)). Analogous to the two modes of understanding, then, we might grant two modes of *meaning*. As we would describe things, since liar sentences do not specify any M-conditions, we will not grant that they are meaningful<sub>1</sub>, though we will, and should, allow that they are meaningful<sub>2</sub>.

It bears noting that, on our pretense account of truth-talk, we might also acknowledge a further sense in which liar sentences are meaningful, though, again, without being meaningful<sub>1</sub>. This has to do with the *content of the pretense* behind the talk, in addition to its props and the principles of generation that apply to them.<sup>7</sup> This content includes the pretense that ‘is true’ is a genuinely descriptive predicate that expresses a property called “truth”. So it might be part of some aspect of the

<sup>6</sup> This allows for a possible contrast between ‘This is a fine red one’ and Chomsky’s (1957) sentence, viz., ‘Colorless green ideas sleep furiously’. The latter sentence is not even understood<sub>2</sub>, if we insist that to understand<sub>2</sub> a sentence, we need how that sentence could be used to make a true assertion.

We should note that, although we are inclined to accept this condition, we need not insist on it, for the points in this paper to go through.

<sup>7</sup> See Woodbridge (2005), p. 171, note 45.

meaning of a liar sentence like (L) that it expresses some content-vehicle's lack of the property of truth. But, on our view, these aspects of meaning would, as it were, all be internal to the pretense and would not constitute any M-conditions that could really obtain or fail to obtain outside the pretense. Hence, they would not render a liar sentence like (L) meaningful<sub>1</sub>, even if they might contribute to our understanding (or: understanding<sub>2</sub>) of that sentence.

### 3 Meaninglessness, generalization and denial

Another challenge our position appears to face pertains to the consequences it might have for the status of the T-schema. The worry is that the meaninglessness we attribute to liar sentences might have the consequence that filling the T-schema in with a liar sentence produces an instance of the schema that is incorrect, and, thus, cannot be accepted. In several publications, Hartry Field (1994, 2001, 2008) has argued that any truth-theorist who does not accept all instances of the T-schema is sunk.<sup>8</sup>

If we do not retain the general correctness of the T-schema (or something close), we will face some notorious problems, for example, that many important 'true'-involving generalizations will be rendered false, unprovable, or worse. A related point Field has made regards the importance of retaining *intersubstitution* in all non-opaque contexts. If we are going to deny the meaningfulness of liar sentences, the question that presses is how we manage to avoid these problems, while still managing to retain these important features.

For our reply, we take a cue from a central element of Field's (1994) pure disquotational account of truth-talk. On his view, the only sentences to which we can legitimately apply the truth-predicate—and, thus, the only sentences that the T-schema applies to—are sentences that we understand, in the sense of 'understand<sub>1</sub>'. Something similar seems to hold for inference-rule deflationism, which sanctions the rules 'true'-in, taking us from a sentence to an ascription of truth to it, and 'true'-out, taking us from a truth-ascription to the sentence to which that predicate is ascribed. The only sentences to which these rules may be applied are the ones that we understand<sub>1</sub>, for only those sentences are *aletheically evaluable*. Although Field (2001) argues, convincingly, that we can *extend* the T-schema to sentences that we do not *yet* understand<sub>1</sub>, the claim that we can do so is, of necessity, constrained: Sentences that cannot be understood<sub>1</sub> are not aletheically evaluable, even if ones that we merely do not presently understand<sub>1</sub>, but could, would be aletheically evaluable.

This has an important consequence for the position we propose, for it means that although there will be incorrect ways of filling in the T-schema, they are not incorrect in the sense of being instances of the schema that fail to be assigned a designated value. They are, rather, incorrect in the sense of not being legitimate

<sup>8</sup> Some discussion at the Seventh Barcelona Workshop on Issues in the Theory of Reference: Paradoxes of Truth and Denotation suggested that this requirement might be a bit strong, but even if that is so, something in the neighborhood will still be required.



instances of the schema at all. So we can accept all of the genuine instances of the T-schema, and even allow for intersubstitution, for any sentence that we do (would, can, or will) understand<sub>1</sub>. Accordingly, this apparent worry about invalidating the T-schema seems to be just that—merely apparent.

One consequence that our view of liar sentences as contentless does have is relevant to recent and other work by so-called “cancellers”, who propose a speech-act solution to the liar paradox according to which we can deny liar sentences non-assertorically, viz., by performing a speech act of denial.<sup>9</sup> While we shall not be concerned with the details behind any canceller’s view, there is something that we can say by way of a general worry about such positions. Cancellers aim to deny liar sentences, where to deny a sentence is to perform a speech act, opposite (or: dual) to affirming. In general, the speech act of denial is used to express *rejection*, where to reject something is to be in a mental state, opposite (or: dual) to accepting. But since rejection is a mental state, what gets rejected is not the sentence one wishes to deny; rather, one rejects *what the sentence says*, or what it *expresses*. If the general denial of a given sentence is intended to express the rejection of what that sentence expresses, then since, on our view, liar sentences lack content, it follows that they do not have anything to reject. So, on our view, cancellers cannot deal with liar sentences by denying them.

#### 4 Semantic characterization and s-defectiveness

Suppose that we are right that we cannot deny liar sentences, with the consequence that we also cannot affirm them, since there is nothing that they express and, hence, nothing to accept or reject. We still face the question of what to say about (L) and other liar sentences; that is, as is widely recognized by those engaged in “liar studies”, we still face the issue of how we will characterize such sentences. And, as is familiar from attempted consistent solutions to the liar, it is at this point that revenge problems generally emerge.

We believe that we can address these issues and avoid the usual problems they appear to generate. Although we shall only have time to sketch a way of dealing with them here, our hope is that what we will provide will be sufficient.

We avoid the “first wave” of revenge problems because we take no positive or negative attitude towards liar sentences, and we neither reason to or from them, or evaluate them semantically—in the sense of ascribing them either a *logical value* or a *truth-value*. On our pretense account of truth-talk, liar sentences do not admit of these sorts of evaluation.

For example, because no M-conditions get specified as prescriptive for the pretenses displayed in a liar sentence like

(L') (L') is false,

there are no M-conditions that would make it correct to utter (L')—that is, no M-conditions that would make it correct to utter the sentence, “(L') is false.” For

<sup>9</sup> Cf. Parsons (1984), Tappenden (1993), Richard (2008).

similar reasons, there are no M-conditions that would make an utterance of “(L’) is not true” correct and also none that would make an utterance of “(L’) is true” correct. Again, this is because no M-conditions prescriptive for the pretenses displayed in any of these evaluations are ever determined, since (L’) itself specifies no M-conditions.

Keeping in mind that liar sentences (and their kin) cannot, in the relevant sense, be understood<sub>1</sub> and, thus, cannot be evaluated in the standard ways, what can we say about them? More directly, how will we characterize them? We propose the following.

As a means for characterizing liar sentences, we introduce a predicate, ‘is semantically defective’ (henceforth, ‘s-defective’), which, for present purposes, is to apply to those sentences, which, while perhaps understood<sub>2</sub>, have no content. More specifically, we are inclined to claim the following, by way of clarifying ‘s-defective’:

- (i) If a sentence, S, is s-defective, then it has nothing, by way of content, which we can accept or reject.

And, as a result,

- (ii) If S is s-defective, then S is not understood<sub>1</sub>.

Moreover,

- (iii) If S fails to specify any M-conditions—either directly or indirectly—then it is appropriate to attribute *s-defectiveness* to S.

Finally,

- (iv) If S is s-defective, then, since S will not be understood<sub>1</sub>, it is not aletheically evaluable,

where, if S is not aletheically evaluable, it cannot (correctly) be assigned (or denied) a truth-value.

To be sure, there is more that we might say about this notion of *s-defectiveness* that we are importing into our vocabulary. But what is crucial here is that ‘s-defective’ applies directly to sentences—actually, to sentence tokens, though the view will not end up looking like a tokenist view, at least in any interesting sense—rather than applying to what a given sentence expresses or applying to a sentence in virtue of applying to what it expresses. Moreover, the expression, ‘s-defective’, applies to sentences that do not possess content, even though such sentences will (or, at least, may) be understood<sub>2</sub>.

By way of shedding further light on the term ‘s-defective’, we might say that for a given sentence, S, S is s-defective under the following condition: the process that would determine what M-conditions S specifies never finishes in the case of S, and, thus, S does not specify any M-conditions at all. To be sure, this does not count as

an *analysis* of the notion of s-defectiveness, as it leaves open the possibility that there are other ways in which a sentence may be deemed s-defective (for example, Strawson's 'This is a fine red one', which includes a demonstrative with no demonstratum), but it will do, for what follows. Let us now apply this approach to liar sentences.

## 5 s-Defectiveness and liars

Consider (L), once again. As we saw, (L) will not specify any M-conditions, which means that, by (iii), (L) will be deemed s-defective, in which case

(5) (L) is s-defective

will be true, and, thus, given the relevant identity,

(6) '(L) is not true' is s-defective

will also be true.

The pressing issue is whether our characterization of (L) as s-defective, and the correctness of ascribing truth to a statement of that characterization, generates revenge problems for us.

Let us turn, then, to a familiar sort of revenge problem that one might pose for our present proposal, as found in

( $\lambda$ ) ( $\lambda$ ) is not true or is s-defective.

In order to see the problem that ( $\lambda$ ) appears to present, suppose that ( $\lambda$ ) is true, false or not true.

For ( $\lambda$ ) to be true it must either be not true or be s-defective—both of which are inconsistent with its being true. It is obvious that inconsistency results, if the left-hand disjunct is true. If ( $\lambda$ ) and the right-hand disjunct are true, then, by (ii), ( $\lambda$ ) is not understood<sub>1</sub>. But if ( $\lambda$ ) is true, then ( $\lambda$ ) is understood<sub>1</sub>. Contradiction. But since a disjunction is true if, and only if, at least one of the disjuncts is true, it follows that ( $\lambda$ ) cannot be evaluated as true.

For ( $\lambda$ ) to be false, it must fail to be s-defective, and it must also fail to be not true. But this means that, if ( $\lambda$ ) were taken to be false, then, via some innocuous reasoning, it would seem to follow that it is true, which, again, issues in inconsistency. Similar consequences arrive, if we hold that ( $\lambda$ ) is not true. After all, if ( $\lambda$ ) is not true then, by disquotation, '( $\lambda$ ) is not true' will be true; hence, it will follow that ( $\lambda$ ) is true, after all. Thus, if ( $\lambda$ ) is not true, it follows that ( $\lambda$ ) is true.

Now, we would characterize ( $\lambda$ ) as s-defective. But if we do, further paradox appears immanent. For if we maintain that ( $\lambda$ ) is s-defective, then we will also accept that '( $\lambda$ ) is s-defective' is true. But now, by disquotation, or-introduction, and enquotation, we seem to be committed to the truth of '( $\lambda$ ) is not true or ( $\lambda$ ) is s-defective', from whence, as we have seen, inconsistency appears to be unavoidable. So, *prima facie*, we too appear to be mired in paradox, having attributed s-defectiveness to ( $\lambda$ ).

To be sure, it certainly *seems* that we are mired in paradox. But, as we will show, we are not. We would be in trouble if we wanted to hold onto classical two-valued logic *and* were to grant that  $(\lambda)$  is a contentful sentence, for then we would seem to be compelled to grant that  $(\lambda)$  is true. But, while we may retain two-valued logic (although we are not committed to this, as we explain in Sect. 7), on our view, paradox is avoided in the case of  $(\lambda)$ , in virtue of the fact that it does not possess any content.

Actually, as we will see, our argument for the claim that, in this case, paradox is avoided, relies on two features, each of which we will motivate. The first is that  $(\lambda)$  is without content. The second is that if a standard, aletheically evaluable sentence is disjoined (or conjoined or otherwise extensionally connected) with a sentence that is without content, then contentfulness cannot be preserved in the resulting complex sentence. We will begin with the first feature, regarding the contentlessness of  $(\lambda)$ . We will then turn to the second.

In order for our attribution of defectiveness to  $(\lambda)$  to generate paradox,  $(\lambda)$  would have to have content, in the sense of specifying M-conditions. But it does not have content, and here is why.

For any content that  $(\lambda)$  would have, both disjuncts are relevant and would have to contribute. This is so because the meaning of a disjunction is a function of the meanings of its parts. So, the meaning—and, thus, the meaningfulness—of  $(\lambda)$  relies, at least in part, on that of its disjuncts. If one of the disjuncts lacks content, then  $(\lambda)$  itself does, too. Accordingly, we will show that  $(\lambda)$  lacks content, by explaining why one of its disjuncts lacks content, where, recall, a given sentence lacks content if it fails to specify M-conditions. In particular, we will show that the left-hand disjunct in  $(\lambda)$ , viz., ‘ $(\lambda)$  is not true’, fails to specify M-conditions and, thus, lacks content.

Now, any M-conditions specified by the left-hand disjunct of  $(\lambda)$  would have to be a product of M-conditions specified by the putative content-vehicle that the disjunct denotes. This is so because, as we mentioned earlier, the principles of generation for ‘is true’ tell us to look to the supposed content-vehicle putatively denoted in an instance of truth-talk and see what M-conditions that content-vehicle specifies. However, the supposed content-vehicle that  $(\lambda)$ ’s left-hand disjunct denotes is just the whole of  $(\lambda)$  itself. This means that what is relevant to determining M-conditions for the left-hand disjunct just is the M-conditions specified by  $(\lambda)$  as a whole. But the left-hand disjunct is part of  $(\lambda)$ , so determining what M-conditions  $(\lambda)$  specifies requires determining the M-conditions that ‘ $(\lambda)$  is not true’ specifies. But that, in turn, requires that we determine what M-conditions  $(\lambda)$  specifies. Accordingly, in order for the left-hand disjunct of  $(\lambda)$  to specify M-conditions, it is required that  $(\lambda)$  already has determined M-conditions. But, of course, M-conditions cannot be settled for  $(\lambda)$  unless, or until, M-conditions are determined for its left-hand disjunct.

So, for any overall M-conditions to get specified by  $(\lambda)$ , there would have to be an impossible sort of semantic bootstrapping, which means that the process for determining what M-conditions  $(\lambda)$  specifies never finishes. Since  $(\lambda)$  fails to specify M-conditions, it follows that the left-hand disjunct does not possess any content, and, so, neither does  $(\lambda)$  itself. As such, both  $(\lambda)$  and its left-hand disjunct lack content. (Notice, though, that both are understood<sub>2</sub>.)

As is evident, our response to the revenge argument relies on a premise, to the effect that only contentful sentences may be disjoined with other contentful sentences to yield a disjunction that is, itself, contentful and thus aletheically evaluable. We will now provide some support for this premise.

Although a conjunction gets its logical value from its conjunctive parts and a disjunction gets its logical value from at least one of its disjunctive parts, both conjunctions and disjunctions get their *content* from *both* of their respective parts. This is so because the content of a complex sentence—viz., a conjunction, disjunction, etc.—is a function from the contents of its parts.<sup>10</sup> But if the content of a complex sentence is a function from that of its parts, then, if any part of a complex sentence lacks content, the same is true of the sentence as a whole. What this means is that the M-conditions for a disjunctive sentence will be a function of the M-conditions for each of its disjuncts. So, if one of the disjuncts of  $(\lambda)$  lacks M-conditions, then  $(\lambda)$  itself lacks M-conditions.

We have thus explained why  $(\lambda)$  specifies no M-conditions and thus has no content. This means that any instance of truth-talk (positive or negative) in which  $(\lambda)$  is the supposed content-vehicle putatively denoted will likewise have no content. Since, as a result, the sentence ' $(\lambda)$  is not true' has no content, disjoining it to another sentence yields a disjunctive string with no content. So, even though ' $(\lambda)$  is s-defective' has content and is true, disjoining this sentence with ' $(\lambda)$  is not true', in order to form  $(\lambda)$  itself, yields a sentence that has no content and, thus, is not aletheically evaluable.

This has an important consequence for the aforementioned revenge argument. Recall that the revenge argument began by allowing that if ' $(\lambda)$  is defective' is true then ' $(\lambda)$  is not true or  $(\lambda)$  is defective' will likewise be true, from which paradox appears to follow. But we have argued that, while we grant that the sentence ' $(\lambda)$  is defective' is true, one cannot disjoin that sentence with ' $(\lambda)$  is not true' and preserve meaningfulness. In this way, we avoid the putative revenge problem on the table, because  $(\lambda)$ 's lack of content blocks the argument that the would-be revenge requires. In the terminology that we favor, because  $(\lambda)$  fails to specify M-conditions, we claim, by (iii), that  $(\lambda)$  is s-defective. Thus, the revenge argument cannot bootstrap  $(\lambda)$  into contentfulness and thereby make it evaluable as true or false (or even as not true or not false).

## 6 Replies to further possible objections: revenge concerns

In this section we consider some further possible objections to our proposed solution. As is well known, a standard aspect of many cases of semantic pathology

<sup>10</sup> An anonymous referee noted that one might object to our assumption of compositionality. The referee is right that one might object to this assumption (one can *always* object to such assumptions). But it bears noting that we are assuming a very *minimal* assumption of compositionality; indeed, one that is generally taken to be uncontroversial, viz., that, in *extensional* contexts, the meaning of a complex formula—in particular, a conjunction, disjunction, etc.—is a function from the meanings of its parts. Since this assumption is generally thought to be uncontroversial and since, anyway, one who resists this assumption will be faced with a rather difficult task, we will not address objections to this (minimal) assumption here.

is their self-attribution of a semantic predicate. Our approach to treating semantic pathology involves the introduction of a new predicate of semantic characterization, so one might question whether cases of self-attribution of our predicate 's-defective' manifest semantic pathology. Another common source of problems, for attempted consistent solutions to the semantic paradoxes, involves appealing to some consequence of the central maneuver the attempted solution makes, as in the standard move to the strengthened liar as a challenge to a basic truth-value gap solution to the simple liar. We consider both a self-attribution worry and a consequence worry, in turn.

## 6.1 Self-application and s-defectiveness

Our solution to the liar paradox involves introducing the predicate 's-defective'. So, a natural question is how we will contend with a case that involves the self-attribution of this predicate, as might occur in a sentence like

(D) (D) is s-defective.

The question, then, is which semantic status we should attribute to (D).

### 6.1.1 *The treatment of (D)*

As is evident, (D) cannot, consistently, be assigned the value 'true'. If (D) is true, then, given what (D) says, it is s-defective. But if it is true, then it is alethically evaluable, in which case, by the contrapositive of principle (iv) regarding s-defectiveness above, (D) is not s-defective. Hence, if we assign truth to (D), it will follow both that (D) s-defective and that it is not the case that (D) s-defective. Thus, (D) will not be assigned the value 'true'.

Suppose, then, that (D) is s-defective. If we were to attribute s-defectiveness to (D), then, given that (D) = '(D) is s-defective', this would entail

(a) '(D) is s-defective' is s-defective.

But, given (a), together with identity, substitution, and both enquotation and disquotation, it would follow that

(b) (D) is s-defective

is true. And, given, again, the relevant identity, substitution and both enquotation and disquotation, this would then seem to entail

(c) (D) is true.

But if (D) is true, then, again, by the contrapositive of principle (iv), it is not s-defective, contrary to what we assumed. Hence, if (D) is deemed s-defective, then it is not the case that (D) is s-defective. Thus, we cannot consistently attribute s-defectiveness to (D).

Suppose, finally, that we assign falsity to (D). If (D) is false then it is not the case that (D) is s-defective. But if (D) is false then, given that (D) = '(D) is s-defective', together with the disquotational schema for falsity, it will follow that it is not the case that (D) is s-defective. Hence, (D) can be assigned the value 'false'.<sup>11</sup>

Before closing this section, it bears noting that we get a parallel argument when we consider

(M) (M) is meaningless.

As is evident, (M) cannot, consistently, be true. But if (M) is deemed meaningless, then, given the relevant identity, together with disquotation and enquotation, it should follow that (M) is not meaningless. That said, (M) can be false, for, if it is false, it is meaningful, in which case it is not the case that it is meaningless.

To be sure, we do not have a *proof* that (D) (or (M)) is false. What we have, and what is all that we need, is reason for thinking that (D) (and (M)) is not paradoxical, as we have reason for thinking that it (like (M)) is false, together with some fairly compelling reasons for thinking that (D) (like (M)) will neither be true nor s-defective (meaningless).<sup>12</sup>

## 6.2 The problem with (R): inconsistency or inexpressibility<sup>13</sup>

We now consider a more worrisome attempt to generate a revenge problem, one involving appeal to a consequence of the central maneuver of our solution to the liar paradox. To get to the more serious challenge, consider principle (iv) again, regarding s-defectiveness,

(iv) If S is s-defective, then...it is not aletheically evaluable.

A consequence of a sentence being s-defective, then, is that it is not truth-evaluable. So we might wonder about the status of a sentence like,

(E) (E) is not aletheically evaluable.

While (E) cannot (consistently) be assigned the value 'true', it can be assigned the value 'false'. But no immediate contradiction also arises if (E) is characterized as not aletheically evaluable instead. At least *prima facie*, then, (E) exhibits a sort of indeterminacy, since it seems that it could be either false or not aletheically

<sup>11</sup> A tokenist (e.g., Goldstein 2009) might try to argue that that the token (D) can also consistently be s-defective (and without yielding expressive incompleteness), which, given the possibility of its falsity, would present a problem of indeterminacy for us regarding (D)'s semantic status. We think that this gambit is unlikely to succeed, given (familiar) revenge problems for tokenists, but we shall not make this case here. For worries about tokenist solutions to the liar paradox, see Armour-Garb (2012b).

<sup>12</sup> Similar (or: dual) reasoning applies in the case of the sentence,

(D') (D') is not s-defective,

which we claim turns out to be true, as does the sentence,

(M') (M') is not meaningless.

<sup>13</sup> We would like to thank Graham Priest both for raising this objection and for discussing the proposed response with Armour-Garb.

evaluable. If so, this would be problematic for our solution to the liar paradox, since the goals of our treatment of semantic pathology generally include avoiding this sort of semantic indeterminacy (familiar from truth-teller sentences, such as, (K) = '(K) is true.').

In fact, however the unsettled status of (E) is merely illusory. This is so because (E) contains a hidden modal which, when made explicit, shows (E) to be false. To see this, notice that to say that a sentence is not alethically evaluable is to say that it cannot be truth-valued, which is to say that it cannot be assigned the value 'true', and it cannot be assigned the value 'false'. But (E) *can* be assigned the value 'false'; *ergo*, (E) can be truth-valued and, thus, is alethically evaluable, contrary to what (E) itself appears to say. This further underwrites assigning it the value 'false'.<sup>14</sup>

While (E) turns out to pose no more problem than (D) (or (M)) does, consideration of this case does point to a more worrisome challenge. Given that we maintain that s-defective sentences are not truth-evaluable, it would then seem to follow further that the truth-predicate does not apply to them. But to see the problem this would yield for our proposal, consider the sentence,

(R) The truth-predicate does not apply to (R).

(R) appears to exhibit many of the traits of (L), and, as with (L), we would argue that (R) is s-defective. (As is evident, we cannot take (R) to be true, false or even untrue, without impending inconsistency.) But if (R) is deemed s-defective, then it would seem to follow that the truth-predicate does not apply to it, in which case, given that (R) = 'The truth predicate does not apply to (R)', it would seem that (R) will be true after all.

If, alternatively, we were to grant that (R) is s-defective but were to refuse to affirm that the truth-predicate does not apply to it, then it would seem that we would fall victim to a charge of expressive incompleteness. Hence, we appear to be stuck with unacceptable inconsistency or intolerable inexpressibility—a clear indicator of pernicious revenge.<sup>15</sup>

<sup>14</sup> Similar reasoning leads us to make the same evaluation of another case, suggested by an anonymous referee:

(E') (E') cannot be (correctly) assigned a truth-value.

No contradiction arises if we assign (E') the value 'false'—if (E') is false, then, assuming double-negation elimination and a little tinkering, it follows that (E') can be correctly assigned a truth-value, which, given what (E') *says*, would make it false. Hence, because (E') can (consistently) be false, it follows that (E') is false. Hence, (E') is false.

A further related case that the anonymous referee suggested is the sentence,  
(E'') (E'') has no alethic evaluation.

While this case differs from both (E) and (E') in eliminating the modal element, like both of them, it still ends up specifying M-conditions (since it suffers from no apparent breakdown on that front), and thus has a correct alethic evaluation. However, since (E'') *says* that it has no alethic evaluation, that makes the value 'false' the correct evaluation for (E'').

Thanks to an anonymous referee, for bringing (E') and (E'') to our attention.

<sup>15</sup> Cf. Armour-Garb (2008).



### 6.2.1 Response

While we appreciate the apparent problem that (R) seems to pose for our proposal, it is important to see why this *appearance* of revenge is, again, illusory. To begin, it is worth pointing out that *application* is a semantic notion—one that is not completely understood. Still, it seems that there are some things that can be said about it, and that particular things *must* be said about it, in order for (R) to appear to raise a problem for our view.

First, let us get clear on how we are supposed to get saddled with a contradiction, from the assumption that (R) is truth-valued. If we attribute truth to (R) then, given what it says, together with the T-schema, it follows that the truth-predicate does not apply to (R). This is not contradictory, but contradiction would seem to emerge, if we were to grant what seems to be a platitude, viz.,

(P1) If a sentence is true, then the truth-predicate applies to it.

After all, if (R) is true, then, by (P1), the truth-predicate applies to (R). But this contradicts the first consequence we drew from (R) being true, namely, that of the truth-predicate not applying to (R).

Suppose, then, that (R) is false. Then it is false that the truth-predicate does not apply to (R), and, so, via some obvious fiddling (and leaving aside any intuitionistic leanings), it follows that the truth-predicate applies to (R). Again, so far, no contradiction emerges. To be sure, given the obvious fiddling, the assumption that (R) is false entails that (R) is false and the truth-predicate applies to (R), and that seems a distinctly weird conjunction to accept. But no contradiction follows, unless we add another principle, e.g.,

(P2) If the truth-predicate applies to a sentence, then the sentence is true.

Given (P2), together with what (R) appears to say, if we assume that (R) is false, we can show that (R) will be true. So, again, we get a contradiction.

Suppose, finally, that (R) is not true. If (R) is not true, then, given the contrapositive of (P2), viz.,

(P2') If a sentence is not true, then the truth-predicate does not apply to it,

together with the fact that (R) = 'The truth-predicate does not apply to (R)' and the T-schema, it will follow that (R) is true. Contradiction again.

Hence, if we grant (P1) and (P2), then it seems that (R) cannot be true, false or, as just shown, not true. But if we do not grant (P1) and (P2), no contradiction follows from any of these alethic evaluations. Hence, it seems that (something like) (P1) and (P2) must be accepted, if this is to pose a problem for our view.<sup>16</sup>

Now, we would not obviously want to reject either (P1) or (P2), given straightforward, non-controversial readings of them. But, as mentioned above,

<sup>16</sup> Without (P1) and (P2), one might conclude that (R) is alethically indeterminate. But, actually, that would not be correct. Rather, without those principles (or ones like them), it would not be clear what the conditions are for the application of the truth-predicate to a sentence, in which case it would be underdetermined—that is, not currently known—what status (R) can or should have.

‘application’ here is a semantic term, so we might be suspicious about its employment or, more generally, its applicability conditions.

Really, the pressing question seems to be this: What is the difference between saying that a sentence is true and saying that the truth-predicate applies to that sentence? One might have thought that those say different things, and that saying that the truth-predicate does not apply to a sentence is more like saying that the sentence is not truth-evaluable than it is like saying that the sentence is not true. However, given (P1) and (P2), this is not actually the case. It seems, rather, that given (P1) and (P2), we are committed to the following two claims (treating the biconditional as material):

(P3) A sentence is true iff the truth-predicate applies to it

and, thus,

(P4) A sentence is not true iff the truth-predicate does not apply to it.

So, if we were to grant (P1) and (P2), we would identify the truth, or the non-truth, of a sentence with the circumstances of ‘true’ applying, or not applying, to that sentence. But since (P1) and (P2) are needed in order to get a paradox going and challenge our proposal, recognizing the identification they yield is the key to seeing the way out of the paradox.

To see why, notice that if a given sentence, *S*, is deemed *s*-defective, then the same will be true of ‘*S* is true’, ‘*S* is false’, and ‘*S* is not true’. But then, given (P3) and (P4), it follows, from our account of *s*-defectiveness, that if a sentence, *S*, is deemed *s*-defective, then one cannot use either the sentence ‘The truth-predicate applies to *S*’ or the sentence ‘The truth-predicate does not apply to *S*’, in order to make an assertion that is meaningful<sub>1</sub>.

So, from attributing *s*-defectiveness to (R), one cannot make a (meaningful<sub>1</sub>) assertion, let alone a true assertion, by uttering ‘The truth-predicate does not apply to (R)’. But there is no “expressive incompleteness” here because there is absolutely no reason for thinking that because (R) is *s*-defective it in some sense follows that we may or should be able to assertorically utter ‘The truth-predicate does not apply to (R)’—any more than that we should be able to assertorically utter ‘(R) is not true’. We can characterize (R) as *s*-defective, and we can take the sentence, ‘(R) is *s*-defective’, to be true. But the sentences ‘(R) is not true’ does not follow from our diagnosis of (R) and neither does the sentence ‘The truth-predicate does not apply to (R)’. These sentences do not follow because, like (R) itself, they are *s*-defective. They cannot be used to make any (meaningful<sub>1</sub>) assertion because, again, like (R), they do not have anything, by way of content, that one can accept or affirm.

## 7 Closing remarks

While avoiding various kinds of revenge problems that thwart other attempted consistent solutions to the liar paradox, our account nevertheless manages to give liar sentences like (L) and ( $\lambda$ ) a sort of semantic characterization, which is different from assigning it a logical value, or attributing to it a truth-value, at least as those

notions are standardly understood. Traditionally, logical values are treated as entities, which are assigned to expressions, as a means for accounting for certain features of the sentences of a given (formal, or logical) language.

We claim that *s*-defective sentences are not alethically or logically evaluable, which is to say that they cannot be assessed either for truth or falsity, or for any logical values. We shall begin by considering alethic evaluability.

In general, we ascribe truth to a sentence when we *accept* what it expresses and we ascribe falsity to a sentence when we *reject* what it expresses. As noted, acceptance and rejection are mental states and are directed at the *contents* of sentences. We can express our acceptance of what a sentence expresses by affirming the sentence, which we can do by asserting that sentence itself, or by asserting a truth-attribution to it. We can express our rejection of what a sentence expresses either by asserting the negation of the sentence or by attributing falsity to it. But for sentences that do not possess any content there is nothing that can be accepted or rejected. Hence, and for the other reasons that we have provided, we cannot (correctly) assertorically attribute either the truth- or the falsity-predicate to them.

Within a logic, we also talk about logical values, which, if we stick with two-valued logic, will be the values, 1 and 0. Now, there are important questions about whether a sentence's having the logical value of 1 or 0 is to be identified with its having the truth-value, true or false. But these are not questions that we can address here. What is important, for present purposes, is that, whether we take the logical values to be 1 and 0, or to be 1, 0, and 1/2, or, perhaps, some other set of values, we maintain that the only sentences that can have any of those values are the alethically evaluable sentences, and, since we also contend that *s*-defective sentences are not alethically evaluable, we conclude that none of them possesses a logical value.

This might make it seem that we are committed to two-valued logic. But we are not. It is compatible with everything that we have said that the appropriate logic to endorse has more than two logical values. It is just that since *s*-defective sentences are not alethically evaluable, they will not be among those sentences that will be assigned a logical value. Indeed, someone concerned, for example, with the indeterminacy found in quantum mechanics may find a reason for assigning 1/2 to certain sentences. But she will still assign 1/2 only to meaningful sentences—only to sentences that are understood<sub>1</sub>.

What this means is that although we will not attribute truth or falsity to any *s*-defective sentence, and we will not assign 1, 0 or 1/2, etc. to any such sentence, the reason for refusing to attribute a truth-value to such sentences is slightly separate from the reason for not assigning those sentences a logical value. We do not attribute truth or falsity to those sentences because, given that they are meaningless, they provide nothing that we can either accept or reject, and, hence, we will neither affirm nor deny them. Moreover, as we demonstrated, previously, any truth-attribution to them will likewise be meaningless. And we do not assign a logical value to them because we take it that only meaningful sentences are alethically evaluable, and that only alethically evaluable sentences can have logical values.

So, our current proposal does not involve taking a view on which logic to endorse. It might then seem, at least *prima facie*, that our focus is on certain features of a natural language. But that, too, would be a mistake. We are *not* claiming that there already is,

in a natural language like English, a predicate that behaves exactly as ‘s-defective’ behaves. So, we would not identify that expression with any of the expressions that we find in a natural language, such as English. Rather, we are *proposing* a predicate, which could serve the task of accommodating expressibility, regarding at least those sentences that lack content in the way that we have explained.

A number of consistent solutions to the liar paradox attempt to unearth or uncover some features of a natural language that had not been adequately, or correctly, recognized. Thus, one finds, for example, that the presence of sentence tokens is taken by some theorists to indicate that there is a means for semantically characterizing liar sentences without falling victim to paradox. Whatever the merits of such accounts, the important point is that those theorists claim to have *discovered* a way of providing a consistent solution to the paradoxes, given only the resources that are already available in, or for, a natural language.

We are *not* trying to do that. Rather than claiming to have found, in a language like English, an expression that can be used consistently to characterize liar sentences, we are proposing a new expression, which people should or could use, in order to describe sentences that, at least on our pretense account of truth-talk, fail to possess any content. While we do think that the pretense-involving nature of truth-talk is an under-appreciated feature of language, we are not claiming to have solved the liar paradox, by discovering and calling attention to this or other under-appreciated features of a natural language like English. In fact, it is completely compatible with everything we have said here that, given certain linguistic demands (e.g., regarding expressibility), our language, or, at least, our current use of that language, results in inconsistency. Understood in this way, one of our aims in this paper was to attempt to satisfy certain expressibility demands without threat of impending inconsistency.

In closing, we should also note that while we accept, and would defend, our pretense view of truth-talk, as we noted previously, the approach to the liar paradox that we have presented is not limited to the assumption of our pretense account, although we think that it fits especially well with it. As far as we can tell, our going approach could be incorporated into other accounts, provided those accounts understand truth-talk to function as a device of content-inheritance, and not as a means of directly specifying M-conditions. Hence, the view is compatible with a number of deflationary accounts of truth.<sup>17</sup>

**Acknowledgments** Thanks to Graham Priest, the audience at the Seventh Barcelona Workshop on Issues in the Theory of Reference: Paradoxes of Truth and Denotation, Barcelona, Spain, June 2011 and an anonymous referee, for helpful comments.

## References

- Armour-Garb, B. (2001). Deflationism and the meaningless strategy. *Analysis*, 61(4), 280–289.
- Armour-Garb, B. (2008). Read and indirect revenge. In S. Rahman, T. Tulenheimo, & E. Genot, Emmanuel (Eds.), *Unity, truth and the liar: The modern relevance of medieval solutions to the liar paradox* (pp. 19–40). Springer.

<sup>17</sup> For a recent discussion of deflationary accounts of truth, see Armour-Garb (2012a, 2012c).

- Armour-Garb, B. (2012a). Challenges to deflationary theories of truth. *Philosophical Compass*.
- Armour-Garb, B. (2012b). No consistent way with paradox. *Analysis*, 72(1), 66–75.
- Armour-Garb, B. (2012c). Deflationism (about theories of truth). *Philosophical Compass*.
- Armour-Garb, B., & Woodbridge, J. (2010). Why deflationists should be pretense theorists (and perhaps already are). In C. Wright & N. Pedersen (Eds.), *New waves in truth* (pp. 59–77). Basingstoke: Palgrave Macmillan.
- Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton de Gruyter.
- Field, H. (1994). Deflationary views of meaning and content. *Mind*, 103, 249–285.
- Field, H. (2001). *Truth and the absence of fact*. Oxford: Clarendon Press.
- Field, H. (2008). *Saving truth from paradox*. Oxford: Oxford University Press.
- Goldstein, L. (2009). A consistent way with paradox. *Philosophical Studies*, 144(3), 377–389.
- Parsons, T. (1984). Assertion, denial and the liar paradox. *Journal of Philosophical Logic*, 13(1984), 137–152.
- Richard, M. (2008). *When truth gives out*. Oxford: Oxford University Press.
- Strawson, P. (1950). On referring. *Mind*, 59, 320–344.
- Tappenden, J. (1993). The liar and sorites paradoxes: Towards a unified treatment. *Journal of Philosophy*, 90, 551–577.
- Woodbridge, J. (2005). Truth as a pretense. In M. Kalderon (Ed.), *Fictionalism in metaphysics* (pp. 134–177). Oxford: Oxford University Press.