

An early version of this paper was published in *Logic and Logical Philosophy* 22 (2013) 445–452.
The present, revised, version appears in Savas L. Tsohatzidis, *Truth, Force, and Knowledge in Language: Essays on Semantic and Pragmatic Topics* (Berlin & Boston: De Gruyter 2020).

Self-reference and the divorce between meaning and truth

Savas L. Tsohatzidis

1. Introduction

Much contemporary work on the semantics of natural languages is guided by the assumption, inspired by Frege, that the meaning of a natural language sentence that is declarative in form is its truth condition. One kind of problem that that assumption (hereafter, “thesis F”) has been claimed to face, and to which no generally accepted solution has so far been proposed, is that there exist declarative sentences that must have the same truth conditions even though they have evidently different meanings. Another kind of problem that it might face would be complementary to the above, and would consist in the existence of declarative sentences that have evidently identical meanings even though they must have different truth conditions. My purpose in this essay is to expose one problem of this second sort, on the basis of inter-linguistic and intra-linguistic evidence concerning the behaviour of a particular kind of self-referential sentence.

A sentence is self-referential, as the term will be here understood, just in case it is the referent of one of its constituents. (Thus, the sentence “The sentence hereby formulated is interesting” is self-referential, since it is itself the referent of its constituent noun phrase “the sentence hereby formulated”.) Liar-sentences (for example, “The sentence hereby formulated is false”) are the most famous self-referential sentences in that sense, but none of those connected with the problem that I am about to describe are liar-sentences and, as I shall argue, the problem’s significance derives in part from the fact that they are not.

2. What the problem is

If any statement of inter-linguistic synonymy is true, the following three certainly are:

- (a) The English noun phrase “the sentence hereby formulated” has the same meaning as the German noun phrase “der hiermit formulierte Satz”.
- (b) The English verb phrase “is an example of an English sentence” has the same meaning as the German verb phrase “ist ein Beispiel eines englischen Satzes”.
- (c) The English verb phrase “is an example of a German sentence” has the same meaning as the German verb phrase “ist ein Beispiel eines deutschen Satzes”.

From these obviously true statements of inter-linguistic synonymy, together with standard assumptions about the compositionality of the processes by which meaningful noun phrases combine with meaningful verb phrases in forming meaningful sentences of English and German, it follows, first, that the sentences in (1) and (2) have the same meaning,

- (1) The sentence hereby formulated is an example of an English sentence.
- (2) Der hiermit formulierte Satz ist ein Beispiel eines englischen Satzes.

and, secondly, that the sentences in (3) and (4) have the same meaning:

- (3) The sentence hereby formulated is an example of a German sentence.
- (4) Der hiermit formulierte Satz ist ein Beispiel eines deutschen Satzes.

It is clear, however, that, although (1) and (2) do have the same *meaning*, they cannot have the same truth conditions: The referent of the noun phrase “the sentence hereby formulated” in (1) cannot be anything other than (1) itself, and (1) is consequently true just in case (1) is an example of an English sentence (which of course it is, hence (1) is true). On the other hand, the referent of the noun phrase “der hiermit formulierte Satz” in (2) cannot be anything other than (2) itself, and (2) is consequently true just in case (2) is an example of an English sentence (which of course it isn’t, hence (2) is false). Similarly, although (3) and (4) have the same *meaning*, they cannot have the same truth conditions. For, the referent of the noun phrase “the sentence hereby formulated” in (3) cannot be anything other than (3) itself, and so (3) is true just in case (3) is an example of a German sentence (which of course it isn’t, hence (3) is false); whereas the referent of the noun phrase “der hiermit formulierte Satz” in (4) cannot be anything other than (4) itself, and so (4) is true just in case (4) is an example of a German sentence (which of course it is, hence (4) is true). It is not the case, then, that whenever two declarative sentences have the same meaning, they have the same truth conditions. And if this is so, a declarative sentence’s meaning cannot be identified with its truth condition, contrary to what thesis F contends.

It might be thought that this sort of problem would be difficult to arise at the intra-linguistic level. In fact, however, it is easy to ascertain its existence at that level, too.

For example, given that the noun phrase “the sentence hereby formulated” cannot be supposed to have a different *meaning* in each one of the different English sentences in which it might occur (though, of course, it will have a different *referent* in each one of the different English sentences in which it might occur), an adequate account of English that aims to compositionally derive the *meanings* of English sentences on the basis of the *meanings* of their constituents (and of the way these constituents are combined) would entail, among other things, that the sentences in (5) and (6) have the same meaning,

- (5) The sentence hereby formulated contains exactly eight words.
- (6) Exactly eight words are contained in the sentence hereby formulated.

that the sentences in (7) and (8) have the same meaning,

- (7) The sentence hereby formulated ends with a verb.
- (8) It is with a verb that the sentence hereby formulated ends.

and that the sentences in (9) and (10) have the same meaning:

- (9) The sentence hereby formulated isn't one that doesn't contain any contractions.
- (10) The sentence hereby formulated is not one that does not contain any contractions.

However, no adequate *truth* theory of English could generate such entailments. Although (5) and (6) do have the same meaning, they cannot have the same truth conditions: in (5), “the sentence hereby formulated” can only refer to (5) itself, and (5) is consequently true just in case it *itself* contains exactly eight words (which it does, hence (5) is true); in (6), on the other hand, “the sentence hereby formulated” can only refer to (6) itself, and (6) is consequently true just in case it *itself* contains exactly eight words (which it doesn't, hence (6) is false). Similarly, although (7) and (8) do have the same meaning, they cannot have the same truth conditions: what “the sentence hereby formulated” refers to in (7) can only be (7) itself, and so (7) is true just in case it *itself* ends with a verb (which it doesn't, hence (7) is false); on the other hand, what “the sentence hereby formulated” refers to in (8) can only be (8) itself, and so (8) is true just in case it *itself* ends with a verb (which it does, hence (8) is true). Finally, although (9) and (10) do have the same meaning, they cannot have the same truth conditions. For, in (9), “the sentence hereby formulated” cannot refer to anything but (9), and so (9) is true just in case it *itself* contains at least one contraction (which it does, hence (9) is true); whereas, in (10), “the sentence hereby formulated” cannot refer to anything but (10), and so (10) is true just in case it *itself* contains at least one contraction (which it doesn't, hence (10) is false). In short, the inter-linguistic evidence falsifies just as clearly as the intra-linguistic evidence does the thesis that whenever two declarative sentences have the same meaning, they have the same truth conditions. And if that thesis is false, it cannot be true, as thesis F contends, that a declarative sentence's meaning is its truth condition.

3. What the problem is not

It might be claimed that the problem that sentences such as those cited above create for thesis F is but an instance of the well-known problem that sentences containing indexical expressions (hereafter, “indexical sentences”) create for thesis F.

That claim does not withstand scrutiny, however. The problem that indexical sentences create for thesis F has two sources: first, that, considered in themselves (i.e. independently of their possible contexts of use), indexical sentences do not have *any* truth conditions; and secondly, that, considered in relation to their possible contexts of use, they can be associated with as many distinct truth conditions as there are relevantly distinct contexts in which they might be used (and so, in effect, with an infinite number of distinct truth conditions). For example, considered in themselves, the indexical sentences “Ich bin Deutsch” and “I am German” do not have any truth conditions; and

when truth conditions are associated with them by reference to their possible utterers, each one of them gets associated with as many distinct truth conditions as there are distinct individuals who might utter it. As a result, if one were to identify, in accordance with thesis F, a sentence's meaning with its truth condition, one would be forced to absurdly conclude, with regard to indexical sentences such as "Ich bin Deutsch" or "I am German", either that they are meaningless or that each one of them has infinitely many distinct meanings. (Incidentally, the standard, Kaplanian, response to *that* problem involves abandoning thesis F. The meaning of an indexical sentence, in the Kaplanian framework, is its so-called "character", and an indexical sentence's character is definitely *not* its truth condition; see Kaplan [1989a, 1989b].)

Notice, however, that neither of the features that make indexical sentences troublesome for thesis F is a feature of the self-referential sentences responsible for the problem under discussion in this essay. For, first, none of these self-referential sentences lacks, considered in itself, a truth condition; and secondly, none of them is susceptible of being associated, when considered in relation to the infinite number of its possible contexts of use, with an infinite number of distinct truth conditions. Rather, each one of them has, considered in itself, a definite, unproblematically specifiable, truth condition; and that condition remains *the same* no matter what the context in which the sentence is used happens to be. So, the problem that thesis F has with these self-referential sentences *cannot* be that it wrongly entails, as it does wrongly entail for indexical ones, either that they are meaningless or that each one of them is associated with an infinite number of meanings. There is, consequently, no basis for supposing that the problem that the self-referential sentences considered above create for thesis F is an instance of the problem that indexical sentences are independently known to create for the same thesis.

One might grant that, for reasons such those just given, indexicality and self-referentiality should be acknowledged to be different things, but nevertheless develop the hope that, since the most famously problematic self-referential sentences are liar-sentences, the problem that the sentences considered in this essay create for thesis F will somehow turn out to be an instance of the problem that liar-sentences create for that thesis.

That hope would be groundless too, however. The problem that liar-sentences create for thesis F is that, although such sentences are meaningful, their meanings cannot be specified, as thesis F disposes one to suppose, by specifying their truth-conditions, since there can be no non-contradictory specification of their truth-conditions. And the source of that problem is not, of course, merely that liar-sentences refer to themselves, but, crucially, that what they *predicate* of themselves is a particular truth-theoretic property, namely the property of being false. (Indeed, it is arguable that self-referentiality not only is not sufficient but may not even be necessary for the generation of the kind of logically paradoxical result associated with liar-sentences; see Yablo [1993].)

Notice, however, that the self-referential sentences discussed in this essay are sharply different from liar-sentences in both of the above respects. First, each one of them has a truth condition whose specification is fully contradiction-free. And secondly, none of them predicates, of what it refers to, a *truth*-theoretic property of any kind (being a sentence of English, being a sentence of German, being a sentence of exactly eight

words, being a sentence that ends with a verb, and being a sentence that contains at least one contraction, are certainly not *truth*-theoretic properties of sentences in any intelligible sense of the term ‘truth-theoretic’). It is therefore groundless to suppose that the problem that these sentences create for thesis F might turn out to be an instance of the problem that liar-sentences create for the same thesis.

4. Conclusion

Since there appears to be no chance of assimilating the problem examined in this essay to already known problems, I submit that it deserves its own distinctive place in the list of problems undermining thesis F.

To be sure, no problem is insurmountable if one is prepared to pretend that the evidence for its existence does not exist, and there are at least two ways in which one might attempt to do that in the present case. The simplest way would be to deny that (1)–(10) are sentences of English or German—in other words, to claim that a grammar of German would be descriptively *inadequate* if it generated (2) or (4), and that a grammar of English would be descriptively *inadequate* if it generated any of (1), (3) or (5)–(10). I take this claim to be too obviously wrong to merit any consideration.

A slightly more complex, though no less futile, way would consist in suggesting that there might be a sense of “meaning” in which the sentences that this essay assumes to be synonymous are not in fact synonymous. In response to such a suggestion, I would simply point out that, although the word “meaning” may have more than one senses, it has *none* in which the sentences in question could be held not to be synonymous. In particular, there is no sense of “meaning” in which the noun phrase “the sentence hereby formulated” acquires a different *meaning* (as opposed to a different referent) each time it appears in a different sentence, and there is no sense of “meaning” in which the noun phrase “the sentence hereby formulated” means anything different from what the noun phrase “der hiermit formulierte Satz” means; given these facts, and assuming that sentence meanings are determined compositionally on the basis of the meanings of sentential constituents, there is no way of avoiding the conclusion that the sentences that this essay assumes to be synonymous are indeed synonymous, despite the fact they *cannot* have the same truth conditions. Trying to ignore the trouble that this fact causes for thesis F by imagining senses of “meaning” that do not exist might be an understandable form of self-deception, but it would be a form of self-deception nonetheless.

References

- Kaplan, David. 1989a. Demonstratives. In Joseph Almog, John Perry & Howard Wettstein (eds.), *Themes from Kaplan*, 481–564. Oxford: Oxford University Press.
- Kaplan, David. 1989b. Afterthoughts. In Joseph Almog, John Perry & Howard Wettstein (eds.), *Themes from Kaplan*, 565–614. Oxford: Oxford University Press.
- Yablo, Stephen. 1993. Paradox without self-reference. *Analysis* 53. 251–252.