Safety, the Preface Paradox and Possible Worlds Semantics

Michael J. Shaffer
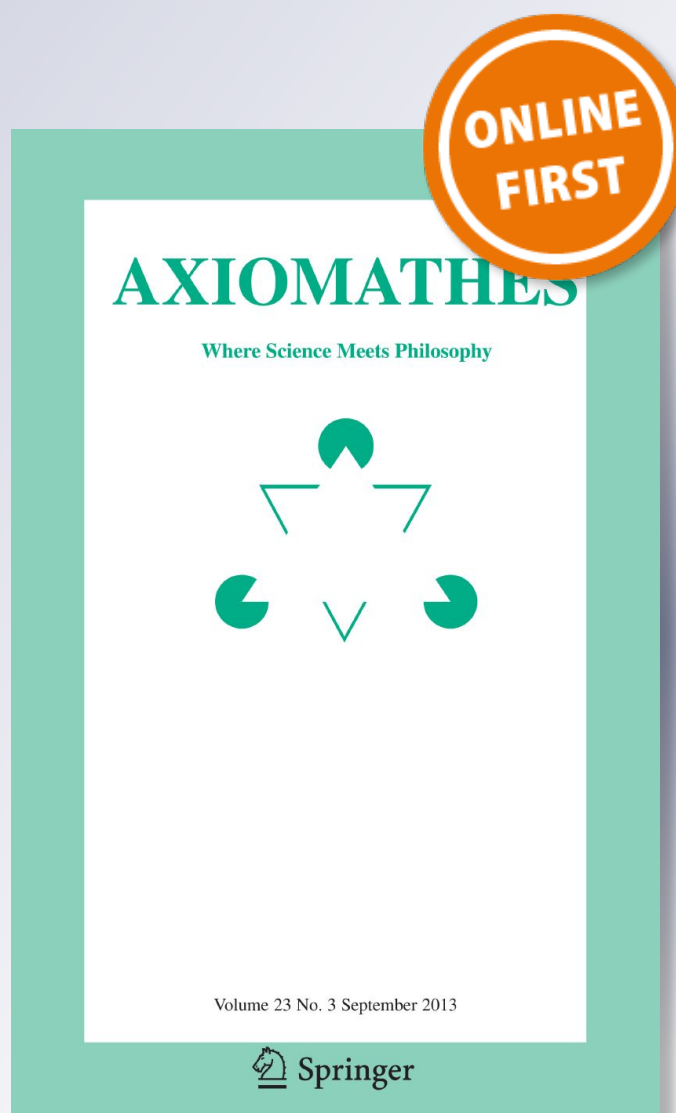
Springer

Springer

ORIGINAL PAPER

CrossMark

Safety, the Preface Paradox and Possible Worlds Semantics

Michael J. Shaffer[1]

Abstract
This paper contains an argument to the effect that possible worlds semantics renders semantic knowledge impossible, no matter what ontological interpretation is given to possible worlds. The essential contention made is that possible worlds semantic knowledge is unsafe and this is shown by a parallel with the preface paradox.

Keywords Preface paradox · Semantic knowledge · Safety · Knowledge · Semantics · Possible worlds

1 Introduction

Possible world semantics holds that the meaning M_P of a well-formed declarative sentence P in a given language L is to be equated with the set of all possible worlds at which that sentence P is true.[1] Where P_{wwf} are the well-formed declarative sentences of language L such that $P \in P_{wwf}$, W_P is the set of worlds $\{w_1, w_2,\ldots, w_n\}$ at which P is true, and W is the set of all possible worlds where for each w_i, $w_i \in W$:

(PWS) The meaning, M_P, of any P in L $= W_P$.

What then is *semantic* knowledge according to PWS? The easy answer is that it is knowledge of the meaning of a sentence and PWS provides an analysis of such meanings. As is standardly the case, knowledge will be understood here to be propositional knowledge and so semantic knowledge will be understood here to be a form of propositional knowledge of the specifically semantic sort. Where SK indicates semantic knowledge of p, A is an epistemic agent and p is a variable that ranges over propositions, this allows for us to introduce the following basic representation of semantic knowledge of p:

---

[1] See Lewis (1970), Cresswell (1978), (1988) and Copeland (2002).

✉ Michael J. Shaffer
    shaffermphil@hotmail.com

[1] Department of Philosophy, CH 365, St. Cloud State University, 720 4th Ave. South, St. Cloud, MN 56301, USA

 Springer

(SKN) $SK_A p$.

We can then see that the defenders of PWS are committed to the following particular analysis of semantic knowledge:

(PWS-SKN) $SK_A P = K_A[W_p]$.

Knowing the meaning of a given sentence P in terms of PWS takes the form $K_A[W_P]$. Here $[W_P]$ is *the proposition* that captures W_P, which, by PWS, is equivalent to the meaning M_P of some P. This is a necessary qualification because one cannot technically know a possible world or set of possible worlds as they are not—except on one particular view—propositions. On this view semantic knowledge of P is then just knowing that each of $w_1, w_2, \ldots, w_n$ is a member of W_P, the set of worlds at which P is true. Such knowledge is knowledge of the extension of $[W_P]$. In any case, PWS-SKN is the expression of the idea that semantic knowledge of a sentence is knowledge of a proposition that captures the set of worlds at which that sentence is true.[2] This is how semantic knowledge is to be understood in terms of PWS and the propositional view of knowledge.

There are, of course, a variety of views concerning the nature of possible worlds, and, hence, a variety of views concerning how we ought ontologically to interpret PWS and so, in turn, interpret PWS-SKN. In this paper it will be shown that some important and commonly accepted interpretations of PWS in terms of the ontology of possible worlds—specifically, linguistic ersatzism, propositionalism, property maximalism and full bore realism about possible worlds—imply that competent speakers of a language could not know the meanings of the sentences that constitute the elements of their linguistic competence given a reasonable understanding of the nature of such knowledge. So, this implies that they could not possibly be competent users of the language(s) in question. This is because the beliefs involved in semantic competence are unsafe so understood. That is to say that these belief do not meet the safety condition on knowledge. However, this conclusion is absurd and that it thereby implicates the versions of PWS-SKN canvassed here if one takes the safety condition seriously. It will be shown that this problem has important parallels with the paradox of the preface that help to illuminate the nature of the problem. Let us begin then by looking at the relevant ontological views of possible worlds and how we are then to interpret PWS given those views of possible worlds.[3]

---

[2] We can see that that this is the proper interpretation of semantic understanding according to possible worlds semantics in terms of Stalnaker's view that possible worlds are maximal properties and we shall examine property maximalism in greater detail in Sect. 2.2. With respect to this view he explains that "…we understand what a particular property is in terms of the range of possible situations in which it would be exemplified (2011, 110)." If worlds are properties and meanings are sets of possible worlds, then it is clear that understandings meanings amounts to understanding the range of possibilities that constitute such meanings. More generally, that this is the correct understanding of the possible worlds semantics account of sentence meaning can be seen in Copeland (2002).

[3] See Melia (2003) on the varieties of interpretations of possible worlds. See Shaffer and Morris (2006) concerning the self-refuting nature of PWS.

2 The Varieties of Possible Worlds Semantics

One attractive and ontologically conservative view is linguistic ersatzism, and this view has been championed by the likes of such notables as Lewis (1947), Carnap (1947), Hintikka (1969) and Montague (1974). There are, however, serious problems with this particular view of possible worlds, especially as it is applied in the analysis of the semantics of actual languages as employed by real linguistic agents. These worries arise in the context of some objectionable epistemic aspects of possible world semantics and linguistic ersatzism that threaten our ability to cognitively grasp the meanings of sentences.[4] But, before we can see what in particular is epistemically objectionable about the ersatzer's version of PWS-SKN and other relevantly similar views, we must consider how the ersatzer conceives of possible worlds. Essentially, on this view a possible world is to be reductively identified with a maximal and complete set of sentences. The elements of any such set are then themselves maximal and consistent set of sentences. So we can define ersatzer possible worlds as follows. Where each S_l is a maximal and consistent set of sentences in some language and $w_j \in W$:

(E) $(\forall w_i)(\exists S_l)\ w_i = S_l$.

Thus, for ersatz semantics the meaning of a sentence P is exhaustively given by linguistically specifying the various ways—in terms of total worlds—that the sentence in question can be true, and this is just to specify the set of maximally complete and consistent sentential state-descriptions in which P is true. We will refer to the set of S_l that constitute W_P for a given P, as S, or the 'S-set' of P. The S-set of a given P, is the semantic, or information, content of P in the Wittgensteinian and Popperian sense that the S-set specifies for P the "…range that it leaves open to the facts (Wittgenstein 1922/1961, 41)".[5] In terms of this view propositional semantic knowledge then must take the following form:

(PWS$_E$-SKN) $SK_A P = K_A[S_p]$.

So, given E and PWS-SKN, semantic knowledge of P is knowing the proposition that captures the S-set of P. Owing to its ontological austerity, ersatzism concerning possible worlds is one of the most well entrenched views concerning the nature of possible worlds. Nevertheless, as we shall soon see, this appealing conjunction of linguistic ersatzism and possible world semantics makes the meanings of ordinary sentences unknowable. Virtually all instances of PWS$_E$-SKN are unsafe. But of course this is absurd, and so either ersatzism or possible world semantics or both must be rejected if one accepts the safety condition. Let us then look at some realist

---

[4] Grasping is intended to connote a state sufficient to entail competence and which allows for various propositional attitudes to be had towards the state in question. As such, the natural interpretation to give to grasping is that it is a form of understanding. See Shaffer and Morris (2010) for more about this issue, especially as it relates to linguistic ersatzism.

[5] Compare Popper (1959, 119–120).

views of possible worlds that identify possible worlds with something more substantial than maximally consistent sets of sentences.

2.1 Propositionalism and Possible Worlds Semantics

Where T_j is a maximal and consistent set of propositions, propositional realism, or propositionalism, takes a world to be defined as follows:

(PR) $(\forall w_i)(\exists T_l)\ w_i = T_l.$[6]

What is then interesting for the purposes of this paper is that, in an important sense, PR is relevantly like E. This is because while those who endorse PR do not reductively identify worlds with maximally consistent sets of sentences, they do identify them with maximal and consistent sets of propositions. Thus, in terms of propositionalist semantics the meaning of a sentence P is exhaustively given by linguistically specifying the various ways—in terms of total worlds—that the sentence in question can be true, and this is just to specify the set of maximally complete and consistent propositional state-descriptions in which P is true. We will refer to the set of T_l that constitute W_P for a given P, as T, or the 'T-set' of P. The analog of PWS_E-SKN for PR is then as follows:

(PWS_{PR}-SKN) $SK_A P = K_A T_p.$

Here we do not need to replace T with [T] for T is itself a proposition according to PR. So, given PR and PWS, semantic knowledge of P is knowing the proposition that is the T-set of P. As we shall soon see, just like in the case of E, this conjunction of propositionalism and possible world semantics makes the meanings of ordinary sentences unknowable. Virtually all instances of PWS_{PR}-SKN are unsafe. But of course this is absurd, and so either PR or PWS or both must be rejected if one accepts the safety condition.

2.2 Property Maximalism and Possible Worlds Semantics

Where U_j is a maximal and consistent property, property maximal realism, or property realism, takes a world to be defined as follows:

(PM) $(\forall w_i)(\exists U_l)\ w_i = U_l.$[7]

Again, as in the case of PR, PM is relevantly like E. This is because while those who endorse PM do not reductively identify worlds with maximally consistent sets of sentences, they do identify them with maximal and consistent sets of properties.

---

[6] See Adams (1974) for an example of this view.

[7] See Stalnaker (1976, 2011) and Forrester (1986) for examples of this view.

Thus, in terms of this view as it applies to possible worlds semantics the meaning of a sentence P is exhaustively given by linguistically specifying the various ways—in terms of total worlds—that the sentence in question can be true, and this is just to specify the set of maximally complete and consistent properties such that P is true. We will refer to the set of U_1 that constitute W_P for a given P, as U, or the 'U-set' of P. But what is known, of course, is the proposition describing such a maximal and consistent property so those who defend PM are also committed to the same basic idea about semantic knowledge that defenders of E and PR are. The analog of PWS_E-SKN for PM is then as follows:

(PWS$_{PM}$-SKN) $SK_A P = K_A[U_P]$.

Given PM and PWS, semantic knowledge of P is knowing the proposition that captures the U-set of P. As in the cases of both E and PR, this conjunction of property maximalism and possible world semantics makes the meanings of ordinary sentences unknowable because virtually all instances of PWS$_{PM}$-SKN are unsafe. But, again, this is absurd, and so either PM or PWS or both must be rejected if one accepts the safety condition.

2.3 Full Bore Realism and Possible Worlds Semantics

Where each V_1 is a maximal and consistent totality of things and $w_j \in W$:

(FBR) $(\forall w_i)(\exists V_1)\ w_i = V_1$.[8]

Thus, for full bore realists the meaning of a sentence P is exhaustively given by linguistically specifying the various ways—in terms of total worlds—that the sentence in question can be true, and this is just to specify the set of maximally complete and consistent sentential state-descriptions in which P is true.[9] We will refer to the set of V_1 that constitute W_P for a given P, as V, or the 'V-set' of P. The analog of PWS$_E$-SKN for FBR is then as follows:

(PWS$_{FBR}$-SKN) $SK_A P = K_A[V_P]$.

---

[8] See Lewis (1986) for the canonical presentation and defense of this view.

[9] One might be argue that the ersatzer need not *identify* meanings with possible worlds understood as sets of sentences and then one might suggest that the ersatzer should simply regard these constructions as useful fictions or models of actual meanings. However, this tactic is not really open to the ersatzer. First, either the ersatzer who accepts possible world semantics identifies meanings with possible worlds, or that view is simply not a version of possible world semantics. Second, without the identification of possible worlds with sets of sentences (maximal or otherwise), the ersatzer loses his metaphysical motivation. If such a theorist holds that the sets of sentences that characterize possible world semantics are merely useful fictions or models then the *ersatzer* cannot simultaneously reject realism about possible worlds while retaining realism about meaning.

Given FBR and PWS, semantic knowledge of *P* is knowing the proposition that captures the V-set of *P*. But, once again, this conjunction of full bore realism and possible world semantics makes the meanings of ordinary sentences unknowable because virtually all instances of PWS_{FBR}-SKN are unsafe. But, again, this is absurd, and so either FBR or PWS or both must be rejected if one accepts the safety condition. So, as we shall soon see, no matter what particular interpretation one gives to PWS, be it E, PR, PM or FBR, we get the same absurd conclusion. Given these versions of PWS semantic knowledge is impossible. In fact, when framed in this way the problem is an understatement of the scope of the problem and the criticism of possible worlds semantics raised here spells doom for any view of possible worlds semantics that treats possible worlds as maximally complete and consistent entities.

Let us the turn to the details of establishing that this is so.

3 Knowledge of Possibilities, Linguistic Competence and the Preface Paradox

The gist of this problem has to do with linguistic competence and the ability to effectively sort bona fide possibilities from impossibilities. On all the views of possible worlds discussed above possible worlds are complete, maximal and consistent. In virtue of these features possible worlds so understood are exceptionally large and complex structures. But, we also know that we are fallible with respect to complex tasks like verifying the consistency of large sets. So, an agent who is a competent speaker of a language knows the meaning of her utterances in virtue of her linguistic competence, but she also knows that she is likely to be wrong about the consistency of some putative possible world descriptions that, if true, specify in part the meanings of sentences that she understands. Here there is then a clear parallel between such modal knowledge and the infamous preface paradox. Let us then introduce the preface paradox and see how it is similar to the issue at hand.

Makinson (1965) discovered the preface paradox and it arises out of a story of the following sort, although there are some variations in the details.[10] Suppose there is an author of a significantly long non-fiction book and that that author is especially diligent in having carefully attempted to establish the truth of every sentence in the book in question. So, the author claims to know that every individual sentence in the book is true. On this basis the author reasons, by agglomeration, that she knows that the conjunction of every sentence in the book is true. Suppose further, however, that, based on past experience of error involving non-fiction books composed of large sets of sentences, the author knows also that it is overwhelmingly likely that she has made a mistake somewhere in the book. So, the author knows that at least one of the sentences in the book is false. She knows that the disjunction of the denials of every sentence in the book is true. As a result, the author is supposed to know both that every sentence in the book is true and that not every sentence in the book is true. But, this result is paradoxically contradictory. The paradoxicality arises chiefly out

---

[10] See Olin (2003, ch. 4).

of the recognition that large-scale and difficult tasks like verifying the truth of large sets of sentences typically involve errors and it is here that we find an important parallel with the task of knowing the descriptions of possible worlds because they are also large-scale sets of sentences. As a result, the same sort of fallibility afflicts the semantic knowledge of language users as described by the forms of PWS discussed above as afflicts the author in the preface paradox story. One compelling way to deal with what is going on in preface cases is to appeal to the safety condition on knowledge and to argue that the author in preface case does not, in point of fact, know that the conjunction of every sentence in the book she authored is true. This is supposed to follow from the knowledge of fallibility and safety.

4 Knowledge and Safety

The safety condition on knowledge is a necessary condition for knowing that has been most systematically defended by Williamson (2000), Sosa (1999) and Pritchard (2007, 2008, 2009a, b) and based on an observation made by Kripke (2011) in a 1986 lecture this condition is implied by any theory of knowledge whereby knowledge implies both truth and belief.[11] It is supposed to reflect the basic idea of the sort of reliability associated with bona fide knowledge. The safety condition can be understood simply as follows:

   If A knows that p, then A could not easily have falsely believed that p.

   This relatively non-technical gloss on safety and it can be made more precise as follows:

   (Safety) $(w_i \vDash K_A p) \rightarrow \neg [<w_i> \vDash (B_A p \ \& \ \neg p)]$.

Here '$<w_i>$' is the set of world sufficiently close to w_i and '$B_A p$' represents that A believes that p. So understood, the safety condition is the claim that if A knows that p at w_i, then A does not believe that p when p is false in worlds sufficiently similar to w_i. This regimentation captures the core idea of the safety condition well. What is useful here is the contrapositive of safety:

   (Contrapositive Safety) $[<w_i> \vDash (B_A p \ \& \ \neg p)] \rightarrow \neg(w_i \vDash K_A p)$.

This version of safety essentially is the assertion that if A could easily have falsely believed that p, then A does not know that p. More technically, it is the claim that if in worlds sufficiently similar to w_i A believes that p and p is false, then A does not know that p at w_i. As noted above, safety has independent merit as a condition on knowledge as it reflects a primitive notion of reliability. As noted earlier, what will ultimately be shown here is that if safety is a necessary condition on knowledge, then the versions of PWS discussed above make semantic knowledge impossible. But, since there are good reasons to endorse safety, it then looks like these semantic

---

[11] This observation has been leveraged into an argument in favor of safety in Shaffer (2017).

theories are false. This problem will be approached first by looking at how safety applies to the preface paradox.

5 The Safety Solution to the Preface Paradox

Let then make the presentation of the preface paradox more precise and see how attention to the safety condition both results in a resolution of that paradox and offers insight into the problem associated with PWS raised here. Where b_1, b_2,…, b_n are the sentences that constitute a non-fiction book authored by A and where n is sufficiently large we can generate the preface paradox as follows:

(PP1) for all n, $K_A(b_n)$.[12]

This is simply the claim that A knows that every sentences in the book is true. This amounts to the following claim for a book with n sentences:

(PP2) $K_A b_1$ & $K_A b_2$ &,…, & $K_A b_n$.

By agglomeration this implies:

(PP3) $K_A(b_1$ & b_2 &,…, & $b_n)$.

In other words, A knows that every individual sentence in the book is true and so knows that the conjunction of sentences constituting the book is true. On the other hand, based on good evidence about our fallibility in general and specifically about our fallibility in preface-like cases, we also have the following principle:

(PP4) $K_A(\neg b_1 \lor \neg b_2 \lor,…, \lor \neg b_n)$.

In other words, based on A's past performance with respect to tasks like the one in question, A knows that there is at least one false sentence in the book. But, PP3 and PP4 are contradictory and so we have a paradox. As Olin (2003) points out, the basic nature of a paradox is that it involves a set of propositions Λ each of which is *prima facie* reasonable to endorse, but where (in the context of background knowledge Σ) the set Λ appears to imply a contradiction. So paradoxes are essentially sets of propositions that appear to be individually rationally endorsable but which cannot collectively be endorsed. This can be because the set Λ is itself internally inconsistent or because Λ appears to imply some proposition p and Σ implies $\neg p$. Let us refer to a given set Λ_i as the *paradox constituting propositions* of paradox i. We can then also present paradoxes as deductive arguments where the members $\Lambda_1, \Lambda_2,…., \Lambda_n$ of a given set Λ are the premises and where they either appear to directly imply (p

---

[12] Of course what is really known are the propositions expressed by these sentences. For the purposes of fidelity to typical presentations of the paradox we can ignore this little complication. See Shaffer (2018) for the oroginal presentation of the safety solution to the preface paradox.

& ¬p) or where Λ appears to imply p and Σ implies ¬p. So in this case {PP3, PP4 and safety} is a paradox generated by the preface paradox story and our background theory of knowledge and in order to resolve the paradox one or more of PP3, PP4 and safety has to be given up.

The safety solution to the preface paradox then involves the recognition that we ought to accept safety and PP4 but reject PP3, thus resolving the paradox, and safety helps to explain why it is not the case that $K_A(b_1 \& b_2 \&,..., \& b_n)$ even where the author has been diligent in checking each sentence in the book. This is because A could easily have falsely believed that $(b_1 \& b_2 \&,..., \& b_n)$ where n is large, and so A does not really know that $(b_1 \& b_2 \&,..., \& b_n)$. This is easily seen by noting that there are clearly many conceivable close possible worlds where the author believes $(b_1 \& b_2 \&,..., \& b_n)$ on the basis of her careful and diligent attempts to verify each b_n but where, nevertheless, $(b_1 \& b_2 \&,..., \& b_n)$ is false because one or more of the sentences in the book is false as per PP4. This is simply because we are fallible knowers, especially in the case of complex tasks like verifying the truth of large bodies of sentences. But that means that the author A does not in fact know the conjunction of the set of sentences that constitute the book in question despite her best efforts to verify every sentence individually. The belief that $(b_1 \& b_2 \&,..., \& b_n)$ is unsafe and, again, this will be true for every such preface case where n is sufficiently large.

6 The Unsafety of Possible Worlds Semantics Modal Knowledge

What this lesson about the preface paradox teaches us about PWS semantic knowledge is quite troubling and it should be obvious. On all of the views of possible worlds introduced earlier and given PWS the descriptions of those worlds are structural analogs of books of vast length. Each such world is described by a vast conjunction of sentences. As such, having semantic knowledge of any contingent expression amounts to knowing the truth of the disjunction of many such "books", where each book amounts to a conjunction of propositions that delineate a possible world. Knowing the truth of the descriptions of such worlds amounts then to knowing that the individual propositions constituting each world are jointly true, no matter what ontological gloss we give to $[W_P]$ and thereby to individual worlds. But, as in the preface case, we know that we are highly fallible with respect to such large-scale tasks. By safety then we can see that something very like the same problem arises in the case of possible worlds semantics knowledge as does in the case of the preface paradox, albeit in a slightly less obvious manner. Recognition of our fallibility and acknowledgment of the safety condition implies that we do not and cannot have such semantic knowledge because such beliefs are unsafe. There will always be close possible worlds where we believe a vast conjunction that is supposed to describe a given possible world but where it is believed even though it is false. So, we cannot be sure that any such w_i is really possible and we cannot then have semantic knowledge that involves knowing w_is. We can see this as follows, where M_P is the meaning of P in L:

(SP1) $SK_A P = K_A[W_P]$.

To know M_P, the meaning of some sentence P in L (i.e. to have semantic knowledge of P), is just to "know" W_P and we saw at the outset here that this amounts to knowledge that each of $w_1, w_2,\ldots, w_n$ is a member of W_P, but this must be expressed as a proposition for such knowledge to be propositional. So, knowledge of the meaning of P is knowledge of the extension of W_P expressed as a proposition.[13] On all of views of possible worlds canvassed above semantic knowledge is knowledge of $[W_P]$ and it takes the form of a conjunction of propositions to the effect that each of the n worlds in $\{w_1, w_2,\ldots, w_n\}$ is a member of W_P. This, in turn, is a conjunction of propositions of the following form: $w_n \in W_P$. Each such proposition $[w_n]$ is a proposition claiming that a maximal and consistent state description is a part of the meaning of P. In other words:

(SP2) $[W_P] = [w_1] \,\&\, [w_2] \,\&\ldots, \,\&\, [w_n]$.

So, semantic knowledge so understood takes the following form:

(SP3) $SK_A[W_P] = K_A([w_1] \,\&\, [w_2] \,\&,\ldots, \,\&\, [w_n])$.

This is equivalent to:

(SP4) $SK_A[W_P] = K_A[w_1] \,\&\, K_A[w_2] \,\&,\ldots, \,\&\, K_A[w_n]$.

Recall here that each $[w_1]$ has the form: $w_n \in W_P$. In these expressions each possible world i is captured by a maximally consistent conjunction of n elementary propositions of the form s_{in} as follows:

(SP5) for all i, $w_i = [s_{i1}] \,\&\, [s_{i2}] \,\&,\ldots, \,\&\, [s_{in}]$.

Knowing the meaning of P on any of these views is then knowing that the vast number of massive state-descriptions in question really are maximal and consistent constituent parts of the meaning of P. But, based on our fallibility with respect to such tasks, it is reasonable to claim that we know that:

(SP4) $K_A(\neg[w_1] \lor \neg[w_2] \lor,\ldots, \lor \neg[w_n])$.

As in the case of the preface paradox, SP3 and SP4 then constitute a paradox. However, semantic knowledge so understood is clearly unsafe and the reason such beliefs are not safe is the very same sort of reason as in the case of our knowing preface claims like PP3. There will always be close possible worlds where a given agent's

---

[13] Notice that on PWS such knowledge must be extensional in order for this view to count as a form of possible worlds semantics. If it was intentsional, then semantic knowledge would be knowledge of some intensional entity that comprehends a set of possible worlds. Clearly that view is not PWS.

beliefs about large conjunctions of the form ($[s_{11}]$ & $[s_{12}]$ &,…, & $[s_{1n}]$) being maximal, consistent components of the meaning of some *P* are believed to be true but are really false, even if we have been utterly diligent in our attempts to verify them. So, on this basis, they cannot be known. All such semantic knowledge is unsafe. The real lesson here from the preface paradox is that SP4 is unimpeachable and the fault in the paradox of possible worlds semantics is to be found in SP3 and PWS. Consequently, it looks like we ought to reject possible worlds semantics as an account of semantic knowledge for it renders such semantic knowledge impossible. It simply requires too much of finite agents like us who have severely limited epistemic capacities. Let us refer to this problem as the safety objection.

7 Can Possible Worlds Semantics Semantic Knowledge be Recused from the Safety Objection?

In order to avoid the rather painful implication of the safety objection, defenders of possible worlds semantics might predictable respond in one of the following ways. First, one might argue that our grasp of semantics that appears to be required for semantic competence and other related semantic phenomena is not a sort of knowledge at all and that SKN is just wrongheaded. Second, one might alternatively argue that semantic understanding *is* a form of knowledge and that SKN is basically correct, but object to SKN-PWS. This might involve one or more of the following contentions: (a) that we need only grasp *some* of the total world descriptions that constitute the full meaning of a sentence in order to have semantic competence, (b) that we need only grasp *parts* of the individual world descriptions that constitute each of the possible worlds in order to have semantic competence and (c) that we need only to be able to grasp world descriptions in principle in order to have semantic competence. Of course, the devil is in the details and any such attempt would need to preserve PWS and would need to tie semantic competence to PWS, in addition to avoiding the safety objection.

These objections to the argument presented here are, however, intolerably vague and/or at odds with the basic notions of possible worlds semantics as it is understood by those who have defended this view. The first response (i.e. denying that semantic competence is a form of knowledge) begs the obvious question of just what semantic competence is if it is not a form of knowledge. Moreover, it seems to be simply wrong and would deprive us of the resources necessary to explain semantic phenomena. We clearly have beliefs about the meanings of the sentences we use, some of which are right and some of which are wrong, and it hard to see how we could explain language learning, the correction of semantic mistakes and meaningful communication if semantic competence is not a form of (transmittable) knowledge. Nevertheless, one might follow Putnam in adopting externalism and hold that meanings are not "in the head" and thus reject the sort of cognitive requirement that is built into SKN. Alternatively, one might adopt a pseudo-Quinean eliminativism about meanings and argue on the basis of deflationary concerns that there are no meanings at all.

In response to the Putnam-inspired externalist suggestion, it should be obvious that the anti-realist defenders of possible worlds semantics cannot really appeal to this tactic without conceding those very notions of possible words, while realist defenders of possible worlds semantics who adopt this defense owe us an explanation of semantic competence, communication and language learning that does not require our grasping the meanings of sentences in the usual sense while retaining PWS. So, trying to avoid the safety objection by claiming that meanings exist but are not "in the head" either conflicts with the notion of possible worlds involved or it leaves us with no explanation of basic semantic phenomena in terms of possible worlds semantics. The sentences that constitute the meanings must be somewhere, and they must seemingly be such that they are cognitively accessible to language users if they are to be useful in explaining semantic phenomena. The second, Quinenan, response (i.e. that there are no meanings at all) similarly cannot be coherently adopted by the defenders of possible worlds semantics, for if there are no meanings, then meanings certainly cannot be possible worlds contra PWS. So rejecting SKN does not really seem to be a particularly viable way to avoid the problem posed here. What then of the second sort of strategic defense? Can possible worlds semantics be saved from the safety objection by appealing to small worlds, by appealing to the idea that semantic competence requires only that we grasp some sub-set of the set of worlds that constitutes the full meaning of sentence on this view of semantics or by appealing to the idea that semantic competence requires only an in principle grasp of possible worlds? Let us see if this can be made to work while preserving the basic principles of possible worlds semantics.

In accord with suggestions made by Stalnaker, Hintikka and Barwise and Perry one might immediately respond that there is no reason that the defenders of possible worlds semantics should not be able to appeal to incomplete or small worlds rather than maximal, or complete, worlds.[14] This would reduce the complexity of the members of the semantic sets that are the meanings of sentences according to possible worlds semantics and thus might render them graspable in a way that would preserve SKN and SKN-PWS. The suggestion here then is that we need not treat world descriptions as maximal, or complete, and one might intuitively suppose that this sort of strategy would be desirable in that it might save possible worlds semantics from the safety objection. If the ways things could have been need not be complete descriptions, then perhaps the descriptions of worlds which meanings are to be identified with could be simplified to the degree that they would be epistemically more feasible for us to grasp. But, unfortunately, this approach alone would not render such semantic knowledge safe in a plausible manner. Even if individual small worlds were not complete in the usual sense, they would still presumably need to be very large in order to capture the semantic richness of the meanings of even simple sentences. Thus, it would still be entirely reasonable to suppose that there would still be close possible worlds where they are believed but false, due to such possibilities of error associated with beliefs about large conjunctions, and the safety objection still seems to cut deep.

---

[14] See Hintikka (1983), Barwise and Perry (1983) and Stalnaker (1986).

So, what of the second version of the second strategy for defending possible worlds semantic against the safety objection? As we saw, one might plausibly try to avoid this objection by claiming that linguistic competence requires is a form of knowledge requiring only that one grasp *some* of the possible worlds that are elements of the meaning of a given declarative sentence as per the relevant definitions noted above. This would obviously reduce the magnitude of the semantic sets involved and hence might render semantic knowledge safe. However, it should be clear that this is not true and it involves introducing some deeply problematic vagueness into possible worlds semantics, while also compromising the basic principles of possible worlds semantics. This latter contention is the case because then there is nothing that corresponds to *the* meaning of a sentence. There will be many meanings for any declarative sentence, each constituted by a sub-set of the semantic set associated with a sentences many of which—presumably—would be adequate for understanding. The former contention (i.e. that this response will not rebut the safety objection) is the case because this suggestion does nothing to reduce the vast size of individual possible worlds descriptions and so knowledge of them would still be subject to the safety objection, Nevertheless, there is some intuitive and independent plausibility to this suggestion. If one needs only to grasp *some* of the possible worlds at which a sentence is true in order to possess semantic competence, then the epistemic task required for linguistic competence is of course reduced. The problem, however, with this suggestion is that it is intolerably vague and it constitutes a significant revision of possible worlds semantics.[15] Absent some principled suggestion concerning which and how many worlds must be grasped in order to use a sentence competently this suggestion cannot even be seriously evaluated. Certainly, one must suspect that linguistic competence would require grasping at least a very large number of possible worlds in order to know when uses of a given sentence are appropriate and that it must involve grasping some reasonable number of canonical worlds corresponding to typical usages. However if this is true, then PWS appears to be false or in need of significant revision that would explain either how the meanings of sentences are to be understood as something other than the total set of possible worlds at which a given sentence is true or how grasping only a sub-set of such worlds is sufficient for the grasping of the meaning of a sentence.

Finally, as suggested above, one might predictably try to avoid the safety objection by claiming that semantic competence can be understood in such a way that it does not require the literal grasping of the semantic set associated with a given sentence. One might try to avoid the complexity-driven problem involved in the safety objection by explaining our grasp of semantics as being secured independently of the grasping of total semantic sets of complete worlds, the grasping of total semantic sets of large but not complete worlds and the grasping of partial sets of complete or large but not complete worlds. This might be achieved by offering an independent account of such semantic competence and noting that it requires only that competent language users understand that the possible worlds semantics meanings of sentences can be in principle formulated. This strategy is however intolerably vague

---

[15] Again, see, for example, Stalnaker (2011, 11).

and it seems to be at odds with core principles of possible worlds semantics. Moreover, it appears to conflict with what defenders of this view assert about linguistic understanding. These charges arise in virtue of the observation that this suggestion detaches semantic competence from PWS and ties semantic competence to some unspecified alternative account of our grasping of meanings in violation of PWS-SKN and potentially in violation of PWS itself. If the sort of grasping of meanings necessary for semantic competence is not knowledge of semantic sets understood in terms of PWS, then what is it a grasp of? Certainly, defenders of possible worlds semantics like Stalnaker (2011, 110) are clear about their construal of such semantic competence in terms of PWS and PWS-SKN, but we cannot even consider whether such a revision of possible worlds semantics would avoid the safety objection absent the details of such a view. Nevertheless, the thought involved here seems to be that semantic competence involves grasping some other unspecified entity/entities sufficient for semantic competence and that such semantic competence presupposes only that the semantic sets understood in terms of PWS could in principle be formulated. But, such an account of semantic competence is not an orthodox form of possible worlds semantics (i.e. it holds that meanings are not (merely/partially) sets of possible worlds) and it would seemingly eliminate the role of possible worlds in the explanation of actual semantic phenomena for obvious reasons related to the in principle qualification on which this defense is predicated (i.e. semantic competence would be secured without requiring the actual grasping of semantic sets). Finally, it is worth noting that the in principle qualification on the grasping of the semantic sets involved would also potentially deprive us of the ability to discriminate language users who truly grasp the meaning of a given sentence from those who do not, for both could understand that it is merely possible to formulate such meanings without having any semantic understanding at all and without being able to grasp even one of the possible worlds at which that sentence is true. For these reasons then, it looks like the safety objection is robust and difficult to avoid, especially while retaining the core principles of possible worlds semantics.

References

Adams R (1974) Theories of actuality. Noûs 8:211–231

Barwise J, Perry J (1983) Situations and attitudes. MIT Press, Cambridge

Carnap R (1947) Meaning and necessity. University of Chicago Press, Chicago

Copeland J (2002) The genesis of possible worlds semantics. J Philos Log 31:99–137

Cresswell M (1978) Semantic competence. In: Guenthner F, Guenthner-Reutter M (eds) Meaning and translation. Duckworth, London, pp 9–43

Cresswell M (1988) Semantical essays. Kluwer, Dordrecht

Forrester P (1986) Ways worlds could be. Aust J Philos 64:15–24

Hintikka J (1969) Models for modalities: selected essays. D. Reidel, Dordrecht

Hintikka J (1983) Situations, possible worlds and attitudes. Synthese 54:153–162

Kripke S (2011) Nozik on knowledge. In: Kripke S (ed) Collected papers, vol 1. Oxford University Press, Oxford, pp 162–224

Lewis CI (1947) An analysis of knowledge and valuation. Open Court Publishing, La Salle

Lewis D (1970) General semantics. Synthese 22:18–67

Lewis D (1986) On the plurality of worlds. Blackwell, Oxford

Makinson D (1965) The paradox of the preface. Analysis 25:205–207

Melia J (2003) Modality. Queens-McGill University Press, Montreal

Montague R (1974) Formal philosophy. Yale University Press, New Haven

Olin D (2003) Paradox. Queens-McGill University Press, Montreal

Popper K (1959) The logic of scientific discovery. Harper & Row, New York

Pritchard D (2007) Anti-luck epistemology. Synthese 158:277–298

Pritchard D (2008) Knowledge, luck, and lotteries. In: Hendricks V, Pritchard D (eds) New waves in epistemology. Palgrave Macmillan, London, pp 28–51

Pritchard D (2009a) Safety-based epistemology: whither now? J Philos Res 34:33–45

Pritchard D (2009b) Knowledge. Palgrave Macmillan, London

Shaffer M (2017) An argument for the safety condition. Logos Episteme 8:517–520

Shaffer M (2018) Safety and the preface paradox. Log Episteme 9:215–219

Shaffer M, Morris J (2006) A paradox for possible worlds semantics. Log et Anal 49:307–317

Shaffer M, Morris J (2010) The epistemic inadequacy of ersatzer possible world semantics. Loge et Anal 53:61–76

Sosa E (1999) How to defeat opposition to Moore. Philos Perspect 13:141–154

Stalnaker R (1976) Possible worlds. Noûs 10:65–75

Stalnaker R (1986) Possible worlds and situations. J Philos Log 15:109–123

Stalnaker R (2011) Mere possibilities. Princeton University Press, Princeton

Williamson T (2000) Knowledge and its limits. Oxford University Press, Oxford

Wittgenstein L (1922/1961) Tractatus logico-philosophicus. Routledge, London