Heidi Savage
May 2019

The Truth and Nothing but the Truth: Non-Literalism and The Habits of Sherlock Holmes

Abstract: Many, if not most philosophers, deny that a sentence like 'Sherlock Holmes smokes' could be true. However, this attitude conflicts with the assignment of true to that sentence by natural language speakers. Furthermore, this process of assigning truth values to sentences like 'Sherlock Holes smokes' seems indistinguishable from the process that leads speakers to assign truth values to other sentences, those like 'Bertrand Russell smokes'. I will explore the idea that when speakers assign the value true to the first sentence, they are not mistaken or confused — that we ought to take these assignments at face value. I show how the most popular alternative to this idea is inadequate for explaining various sentences involving fictional names. In addition, I offer evidence that these truth value assignments to sentences are tracking semantic content rather than pragmatic effects. I also motivate a methodological constraint on semantic theorizing that dictates adopting a literalist approach to a certain type of fictional discourse.

1. Introduction

We engage in discourse about characters from works of fiction quite frequently. This discourse seems both meaningful and able to convey truths, not only within the context of theoretical analysis and criticism, involving abstract literary properties, but also within more ordinary contexts, involving simple everyday properties, here called "fictive discourse" to distinguish it from the former kind. For instance, sentences such as 'Sherlock Holmes smokes', 'Heathcliff is a tyrant', and 'Elizabeth Bennet has a sharp wit' — drawn from the fictional works respectively written by Doyle (2002), Bronte (1976), and Austen (1993) — all seem to express something true. But, we also think in saying something true, we are talking about something — an existent individual or an actual state of affairs. It follows that, if the previous sentences are true, then they must be about an existent individual or an actual state of affairs. But no one believes that there is any pipe-smoking Sherlock Holmes, tyrannical Heathcliff, or sadly, any witty Elizabeth Bennet. In other words, we know that the previous sentences occurring within the context of fictive discourse are not about anything,

and yet our intuition — that the previous sentences are true — remains. Clearly, fictive discourse raises a puzzle in need of some kind of resolution.

The previous facts generate what I will call the "puzzle of fictive discourse," a challenge to almost any widely accepted theory of meaning, including Fregeanism, Russellianism, Davidsonian, Possible Worlds Semantics, and their variants, and which can all be broadly characterized as truth conditional in nature.[1] Truth conditional theories, for the most part, hold that the meanings of declarative sentences consist in their informative, or truth-apt content. Furthermore, these theories also commonly adopt the previous principle that, in to order to utters truths, there must be some individual or state of affairs about which we can utter them, as a semantic principle for evaluating the previous types of sentences — simple subject-predicate sentences. This semantic evaluation principle can be expressed more formally as the principle that, for any simple subject-predicate sentence, it is true just in case the individual corresponding to its subject has the property corresponding to its predicate — what I will call the "traditional" analysis of simple subject-predicate sentences.

The traditional analysis clearly dictates that if there is no Sherlock Holmes, Heathcliff, or Elizabeth Bennet, then any sentence that purports to be about any of these individuals, must be literally false. On this way of thinking, our sense that these sentences are true must be explained by some kind of non-literal information they somehow convey. One way to deal with the puzzle, then, is to reject our intuition that the previous sentences express anything true at all — a "non-literalist" approach to the puzzle of fictive discourse (see Currie, 1988; Evans, 1982; Lewis, 1983; and Walton 2006, among many others). In

[1] For the respective original proponents of these views, see Frege 1956; Russell 1919; Davidson, 1967; and Carnap, 1947. For more modern proponents of each respective view see Schiffer, 2000; Braun, 2003; Kolbel, 2001; and Lewis, 1986.

contrast, a “literalist” approach to the puzzle, the approach I support, takes seriously the idea that the previous sentences do express something true, and that they do so as a matter of their semantic content (see Deutsch, 2001; Elgin, 1986; Ludlow, 2006; Tiedke, 2011; and Martin and Schotch, 1974).[2] Adopting a literalist position, then, clearly conflicts with the traditional analysis — a well-accepted, well-warranted account.[3]

So, then, why defend a literalist approach if it means giving up on such an appealing account? I offer three arguments for doing so. The first argument is that the non-literalist's favored explanation for the appearance of the truth of certain sentences from fictive discourse — the story operator account — fails to apply to certain types of fictive discourse. And it fails because these types of fictive discourse are not plausibly analyzed using a story operator account, thereby reviving literalism as a potentially viable approach. The second argument offers positive support for adopting literalism based on the fact that the assignment of the value true to a sentence like (1), for instance, remains constant even in a context that demands a literal interpretation of that sentence. The third argument shows that literalism, in spite of initial impressions, can ultimately respect truth conditional theories of meaning better than non-literalism can.

In the next section, I describe the puzzle of fictive discourse in more explicit terms, as well elaborating on both non-literalist and literalist approaches to it. In sections three through five, I develop and explain the three main arguments in favor of literalism, spending a significant amount of time in section three on the most complex of all three — the

---

[2] Because non-literalist and literalist views differ significantly from one another, and I am interested in the general approach itself, I offer a characterization of these that aims at neutrality with respect to which of them is correct. The same is true of my characterization of truth conditional theories of meaning as well.

[3] I defend a specific literalist analysis of the content fictive discourse in my previous work where I reject a traditional analysis of its content (2011).

argument that non-literalism fails. Finally, in section six, I consider objections.

2. The Puzzle of Fictive Discourse

To illustrate the nature of the puzzle of fictive discourse more clearly, consider the following two pairs of simple subject-predicate sentences:

(1) Sherlock Holmes smokes
(2) Sherlock Holmes wears a tutu

(3) Bertrand Russell smokes
(4) Bertrand Russell wears a tutu

Now consider their intuitive truth values. Respectively, true then false for each pair. The fact that speakers share these intuitions is empirically supported by a study conducted by Piccinini and Scott (2010) reporting that over 80% of informed speakers will assign the value true to sentences like (1), and false to sentences like (2). Likewise, we should expect that informed speakers will repeat this pattern of assignments for sentences (3) and (4). Furthermore, the assignments of truth values to sentences like (1) and (2) occur as spontaneously, and as effortlessly, as they do for sentences like (3) and (4). For Piccinini and Scott, this fact suggests that the procedure for assigning these truth values are the same in both cases, and that therefore our treatment of both pairs should be uniform.[4]

On the issue of uniformity, in general, semanticists seek to provide accounts of the meanings of sentences that are systematic. To be specific, they seek accounts that are compositional in nature, which involves identifying a sentence's type of basic syntactic parts, their associated meanings, and then specifying rules for how combine those meaningful syntactic parts to form more complex types of meaningful phrases. The

[4] Although Piccinini and Scott's examples involved claims about Santa Claus and Mickey Mouse, I see no reason not to generalize their results to sentences like (1) and (2).

traditional analysis of the content of a simple subject-predicate sentence is a good example of a compositional semantic rule. Since sentences (1)-(4) do not apparently differ with respect to their basic syntactic parts, or their mode of combination, they should all get the same semantic analysis.

There are, then, two apparent facts that must be explained by any semanticist: that speakers' assign the truth values they do to sentences (1)-(4); and that these sentences must be given the same semantic analysis. These two apparent facts along with two other apparent facts — that sentences (1) and (2) lack subjects that have referents, and that the meaning of a simple-subject predicate sentence is given by the traditional analysis — combine to produce the puzzle of fictive discourse.

The puzzle is that not all of these apparent facts can be facts that we take at face value. One of them must be rejected. I will not explore the idea that we can give up compositionality, or that sentences (1) and (2) lack referents. First, the reason for not exploring the idea of giving up compositionality is simple. It threatens the entire point of even attempting to give any analysis of any natural language. Second, the reason for not exploring the idea of positing a referent for sentences like (1) and (2) is that any referent we could posit would not help us make sense of the assignment of truth values to sentences (1)-(4).[5]

---

[5] Many, including, Kripke (1973 ms.), VanIwagen (1977), and Zalta (1983), among others, do deny the assumption that fictional names lack referents. Instead, their referents are abstract objects — fictional characters. This approach, however, is motivated by different considerations from those considered here. It aims at solving the problem of how sentences like 'There are more characters in *Pride and Prejudice* than in *Wuthering Heights*' can be true*,* which presents a different problem than the puzzle of fictive discourse, what I call the "problem of meta-fictive discourse."

The only options left, then, are to either reject the idea that the assignments of values to sentences (1) and (2) reflect anything about the semantic content of those sentences, or to reject the traditional analysis of their semantic content. It seems that we are faced with a dilemma. If fictional names have no referent, then the traditional analysis entails that fictive sentences like those considered must be false. If so, we cannot take the the tendency of speakers to assign the value true to some of them at face value. In contrast, if we do wish to take these assignments at face value, then we cannot maintain the traditional analysis.[6] Defending either side of the dilemma entails certain challenges.

The non-literalist accepts the traditional analysis. Therefore, they must explain the fact that speakers assign different truth values to sentences like (1) and (2), since their view entails that both are equally false.[7] Literalists, by contrast, although they accept the truth value assigned to sentences like (1) at face value, must then face the challenge of offering a semantic analysis of sentences (1)-(4) alternative to the traditional analysis, a significant burden, given its difficulty, and the traditional analysis' plausibility.[8]

Having explained the puzzle of fictive discourse, and having surveyed two main positions with respect to it, I will now focus on arguments in favor of the approach I believe

---

[6] There is a third position on the truth value of sentences like (1) and (2). This position denies that these sentences have any truth value at all. They are instead indeterminate. While I am not opposed to this idea, at least not on he grounds of anything I am committed to here, I will not discuss this position further for two reasons: first, inserting the required qualifications into the main discussion would compromise accessibility, but second, because speakers do assign a sentence like (1) the value true, this position must address the same issue as a non-literalist. For more details about the problems associated with this position, see Braun (2003).

[7] For explanations of this fact, see Bertolet, 1974; Braun, 2003; Evans, 1982; Lewis, 1978; and Walton, 2006.

[8] In addition to my own analysis offered elsewhere, there are also others who offer such an analysis, see Deutsch, 2000; Elgin, 1983; Ludlow, 2006; Martin and Schotch, 1974; and Tiedke, 2011.

is correct — literalism. I begin by showing how literalism's main competitor — the non-literalist story operator theory — fails.

3. The Non-literalist Story Operator Account

The most popular non-literalist explanation for the truth values that speakers assign to sentences (1) and (2) is the story operator account (Lewis, 1978). This approach to the puzzle holds that sentences like (1) and (2) appear have truth values only because they are associated with other sentences that are true. The sentences that are versions of the previous sentences qualified with a story operator, such as 'In the story' or 'According to the story'. So sentence (1) seems true because sentence

(1') In the story, Sherlock Holmes smokes

is true. And sentence (2) seems false because sentence

(2') In the story, Sherlock Holmes wears a tutu

is false. This view, then, explains the difference between (1) and (2) in virtue of the fact that (1) expresses something true when qualified with a story operator, while (2) does not.

For the non-literalist, the semantic content of sentences (1)-(4) is still given by the traditional analysis. However, speakers can still communicate something true, because these sentences are pragmatically associated with other sentences, which are qualified with a story operator, and these sentences can be true (Bertolet, 1984).[9]

I now turn to discussing some concrete samples of fictive discourse, beginning with a case that appears to vindicate the story operator account, in order to give the account its

[9] This is, however, only one way to understand the story operator account. We could take the story operator analysis as yet an additional literal reading of sentence (1). That is, sentences like (1) are ambiguous between traditional and story operator readings. This would not be a non-literalist account, however, and this is the position I am critiquing. For that reason, I will not pursues this account any further here..

due. I then discuss two other conversations that reveal the shortcomings of the story operation account drawn, respectively, from fan fiction, and fictive modal discourse.

3.1 Vindicating the Story Operator Account

Imagine that two individuals, Adam and Sam, are having a conversation about the habits of Sherlock Holmes, in which the background assumptions in play are those of most ordinary speakers when they assign the value true to a sentence like (1). Now imagine that Adam begins the conversation by uttering a sentence that, according to the non-literalist, must be false. We will also imagine that Sam is not in any mood for charitable pragmatic re-interpretation.

Conversation 1

> Adam: Sherlock Holmes was a great detective who mainlined cocaine.
> Sam: Sherlock Holmes couldn't have done that, or anything else for that matter, since he doesn't exist.
> Adam: Well, OK, but at least according to the stories, he was a great detective who mainlined cocaine.
> Sam: Of course, that's true.

Conversation 1 patterns exactly as the non-literalist story operator theorist predicts fictive discourse should proceed. We have an initial literal interpretation of Adam's first utterance by Sam. We then see Adam make an explicit appeal to a story operator in order to clarify what is being said, and we see that Sam is then prepared to accept that that was what was said, and that it is true. Of course, given Sam's familiarity with the stories, interpreting the initial utterance as having a traditional reading is a rather churlish act on his part. Still, this does not show that the traditional reading is incorrect. If the traditional reading is the initial reading, and it is read as false, and subsequent qualifications with story operators shifts that evaluation to true, then this is a case that vindicates the story operator account.

3.2 Trouble for the Story Operator Account: Fan Fiction

I will now consider a different conversation that illustrates the story operator account's inability to cope with other forms fictive discourse. Let us assume that Sam maintains the attitude that it is impossible that a person, fictional or otherwise, could use mind altering substances to such an extent and do what Holmes did. Imagine that, on this day, unlike the previous one, Sam is in a rather Straussian mood.

Conversation 2

> Adam: Sherlock Holmes was a great detective who mainlined cocaine.
> Sam: Sherlock Holmes couldn't have been like that. No great detective could be a drug addict.
> Adam: Well, OK, but according to the stories, he was.
> Sam: Well that might be what the story says, but I just don't think that Holmes could have done that, and been such a great detective. Makes me wonder if maybe a it `was really Watson who was the great detective, and Holmes just took all the credit.
> Adam: Well, that's certainly not what's in the stories. But either way, I don't think Watson had the brain power to pull it off.

It is these kinds of musings that lead to what is now a thriving and vast body of literature known as "fan fiction" — discourse based on an original work of fiction, but that transgresses what is true according to that work. We can understand the disagreement between Adam and Sam, then, as a disagreement within the realm of fan fiction, in which Adam represents the conservative position that we cannot take such licenses with a work, and Sam represents the more progressive idea that we can do so.

Now what does conversation 2 look like according to the story operator account? Adam's first utterance would be literally false, but true if qualified with a story operator, which is made clear by his third utterance. If the story operator theorist were correct, then the issue of the habits of Sherlock Holmes ought to be settled. Yet the conversation takes a different turn, not predicted by the story operator account, leading to questions about what

might really or plausibly be true about Sherlock Holmes, independent of what is in the story itself. The fact that the qualified interpretations of Adam's utterances do not settle the issue about what is true of Sherlock Holmes, shows that Sam is assuming that not only are there facts about what is true according to a story, there is another kind of interpretation of fictive discourse concerning what is really true of the characters from works of fiction, outside of the context of a story.

Now however we value such conversations, they occur fairly often, and at least to those involved, seem perfectly reasonable — a fact to which any avid reader or writer of fan fiction could attest. A complete account of fictive discourse, therefore, needs to address these conversations, simply because they exist. It is difficult to see how the story operator approach could explain them. Fan fiction is a kind of of fictive discourse that contains sentences truth values are clearly not determined by what is true or false according to a story.

The story operator account, then, fails to exhaust all forms of fictive discourse. This opens up room to once again seriously consider a literalist position. In addition, as we'll now see, fan fiction is not the only form of fictive discourse that contains sentences whose truth or falsity does not depend upon what is true or false according to a story. Fictive modal discourse also displays this characteristic to an even stronger degree, to which I will now turn to discussing.

3.3 More Trouble for the Story Operator Account: Modal Discourse About Fiction

As I will argue, claims from fictive modal discourse can be interpreted in two separate, but equally legitimate ways — in some literalist way inconsistent with the traditional analysis, or in the qualified story operator way.

To begin with, consider this hypothetical conversation between Adam and Sam concerning whether Sherlock Holmes could have failed to have been a detective — a bit of fictive modal discourse. Suppose Adam and Sam just finished taking a course in the philosophy of modality:

Conversation 3

> Adam: So, what do you think? Could Holmes have been a criminal? Could he have been something other than a detective?
> Sam: Well, of course, according to the story, Holmes has the same modal properties as any other ordinary person, and their careers are not essential to them. So, yes, he could very well have been something other than a detective.
> Adam: But Sam, don't you think there is also some sense in which Holmes would not be Holmes if he was not a detective? I mean, he has no properties outside of those decided upon by Doyle, and those decisions make Holmes what he is. Without the Doyle's act of penning the stories, Holmes would be nothing at all.

As we can see, in conversation 3, there are two different interpretations in play when discussing the modal properties of Sherlock Holmes — the story operator interpretation, and another interpretation concerned with Sherlock Holmes's actual modal properties.

On the story operator interpretation, the answer to Adam's first question is obviously "yes," as Sam points out. But, as Adam also points out, there is another intuition about the appropriate way to answer this question — that Sherlock Holmes could not have been other than a detective. But why should we think that such claims are true, and what would make them true if not what is true according to the story?

Adam's explanation, in conversation 3, is somewhat convoluted. However, we can still extract a certain simple kind of explanation for how and why we might think that Holmes is essentially a detective. Consider this idea: Doyle created Sherlock Holmes by stipulating his properties, and if his properties are stipulatively true of him, then they are necessarily true of him. This simply explanation, it turns out, is too simple, however. It entails that if

Sherlock Holmes is essentially a detective for the previous reason, then all of the properties he has, equally the product of Doyle's stipulations, will also be essential to him — like the particular color of socks he wears, for instance. Now the idea that Sherlock Holmes is essentially a detective is reasonable, but surely that it is essential that he wears a certain color of socks is not.

A more complex explanation distinguishes between what makes Sherlock Holmes a member of the kind fictional entity from what makes him the character that he is. What makes Sherlock Holmes fictional is indeed the fact that he was created by an act(s) of stipulation in which a group of properties became associated with the name 'Sherlock Holmes'. However, this act determines does not determine Sherlock Holmes' essential properties. Instead, it determines only that he is a member of the kind fictional entity. So what does determine his essential properties? In order to answer this question, I focus now not on what makes Sherlock Holmes fictional, but on what makes him a character.

The answer involves appealing to the acts of stipulation that created Sherlock Holmes in the first place, but only to a certain subset of them. This subset of stipulative actions will be those that determine Sherlock Holmes' character — his essential properties. I take it that the stipulations that will define a fictional character's essential properties are those that the author chooses in order to fix that character's role in the story, not merely those whose coalescence are sufficient to form a fictional character, but those are definitive of that fictional character's role within the events of the story. Surely, being a detective is part of Sherlock Holmes's role in the story, as is being exceedingly smart. It follows then that Sherlock Holmes is essentially a detective, a brilliant one at that, among other things. Having offered some explanation for thinking that our modal intuitions about fictional

characters are justified, I will now turn back to their impact on the story operator account.

The problem that these modal intuitions present for the story operator account is it simply does not seem that they can be accommodated. It seems that it could never come out true on the story operator account that Holmes is essentially a detective. And this is because, on that view, the only way that Holmes could essentially be a detective would be if it were true according to the stories that he was, which Sam points out is false.

Fictive modal discourse also challenges the traditional analysis, which entails the falsity of literalism, more directly. Consider the fact that when asking speakers to evaluate claims like those above, they are sometimes flummoxed. But this is not due to the fact that they are confused about whether to evaluate these claims by relying on a story operator account or the traditional analysis — the two non-literalist options. If this were the explanation, then their confusion would be fairly odd, since the answers to questions about the modal properties fictional characters have according to the stories, or according to the traditional analysis, are pretty obvious.

So, then, what are speakers confused about then when asked to evaluate fictive modal claims about fictional characters? The only answer that seems right is that they must be confused about which set of modal properties they are being asked about — those that a fictional character actually has, or those that they have within a story. The false traditional reading — what I think of as the “philosophers’ reading” — doesn’t even typically occur to speakers, not even, in my experience, to philosophers themselves. This shows that, at least concerning fictive modal discourse, the traditional analysis has no role to play in analyzing

its content.[10] Fictive modal discourse, then, presents both indirect challenge to the traditional analysis — in virtue of the failure of the story operator account — and also a more direct challenge that illustrates its complete irrelevance with respect to evaluating fictive modal claims.

While the simplicity and plausibility of the story operator account as a solution to the puzzle of fictive discourse is to be admired, as we've seen, it cannot explain certain cases of this kind of discourse. Specifically, it fails to explain fan fiction, as well as fictive modal discourse. This offers even more reason to consider accepting the idea that some sentences from fictive discourse are just plain literally true. To justify taking up this literalist approach to fictive discourse, I will now offer two positive arguments in favor of adopting it.

4. Testing the Literalness of Fictive Discourse

One positive reason for adopting a literalist position comes from considering the results of a test I devised at some point, and that I call the "neutral context" test. This test is used to isolate a sentence's semantic content from any pragmatic content it might be associated with. This is done by asking speakers to assign a truth value to a test sentence, which is presented without any background information, or at least, with very little background information. In fact, the original presentation of sentences (1)-(4) was intended to be an instance of presenting sentences in a paradigm neutral context, in which a speaker merely mentions, rather than uses, the test sentence in question. This might also occur in a context in which a speaker simply displays sentences on a chalkboard, and then asks speakers to evaluate them for truth. The idea is to refrain from embedding a sentence within a particular

[10] Note that this particular argument applies equally to the ambiguity version of the story operator account as well, given that this account also predicts only two interpretations of sentences from fictive discourse — the story operator and the traditional analysis, and these both fail to account of modal fictive discourse.

conversation, so that the pragmatic effects involved in acts of communication are unlikely to be triggered. If a test sentence is presented in a context in which no pragmatic effects are triggered, then the only available interpretation of that sentence will be its literal interpretation. A given hearer's assignments of truth values to test sentences could then be trusted to be tracking semantic rather than pragmatic content. If the presentation of sentences (1)-(4), then, is in fact an instance of presenting sentences in a paradigm neutral context, it follows that we ought to the truth value assignments to those sentences at face value. In other words, we should adopt the literalist position on the puzzle of fictive discourse.

One issue to inquire about is whether the neutral context test really does isolate a sentence's semantic content, whether it is valid in the scientific sense, measuring what it purports to measure. To show that the test is valid, I will now consider a sentence that is almost never interpreted literally. However, it is plausible to think that it would be so interpreted that way in a neutral context. The upshot is that if an entrenched pragmatic interpretation can be dislodged by the neutral context test, this illustrates its scientific validity. The sentence 'I am not going to die' serves as a particularly good illustration. Rarely do speakers and hearers ever interpret this sentence as conveying the idea that the utterer of the sentence is immortal — its actual content. Instead, it almost always gets uttered and interpreted in conversational contexts in which the speaker desires to express the fact that they are not going to die anytime soon — as a result of a good prognosis from a doctor, for instance. It is therefore an instance of a sentence with an entrenched pragmatic interpretation. But now consider whether this interpretation will survive the neutral context test. Consider how speakers might interpret the sentence 'I am not going to die' if it is

merely displayed outside of any particular conversation. I think it is likely that the pragmatic interpretation will not survive the test. Admittedly, I have not yet run any experiments to confirm this hypothesis, but I would place my bets on the outcome that speakers would in fact interpret an utterance of this sentence as false in a neutral a context. The natural interpretation would most likely be the immortality interpretation — its literal content. If the neutral context test can induce a literal interpretation even in cases like the one considered — those in which the sentence is almost never interpreted literally — it has proven itself as a particularly sensitive instrument for detecting differences in semantic and pragmatic content.

While the neutral context test gives us some relatively strong evidence in favor of a literalist approach to the puzzle of fictive discourse, it is nevertheless subject to the same vicissitudes that affect the validity of any scientific instrument of measurement. For this reason, I now offer a more conceptually oriented reason for adopting literalism, having to do with a specific methodological constraint that governs the relationship between theory and data. Applying this methodological constraint to truth conditional theories again justifies adopting literalism about fictive discourse.

5. The Methodology of Constructing Theories of Meaning

While non-literalists may face challenges, nevertheless, they can maintain a commitment to the traditional analysis, which is a plausible, entrenched, well-accepted, and well-warranted analysis of the content of sentences like (1)-(4). It is because non-literalism respects the traditional analysis of the content of simple subject-predicate sentences, that it also appears to be able to more readily respect the widely accepted truth conditional approach to theories of meaning — its best justification.

However, I will now argue that, given a certain plausible methodological constraint on theories of meaning, truth conditionalism itself requires rejecting the traditional analysis, thereby undermining the entire reason for adopting a non-literalist approach in the first place. The methodological constraint, involves a coherence requirement on any theory of meaning. It is based on the idea that any semantic theory of natural language ought to fit with speakers' linguistic behaviors, and that if it fails to do so, we ought to reject it, or engage in its revision. As applied to truth conditional theories, it is reasonable to suppose that this constraint requires that the correct analysis of the content of simple subject-predicate sentences, like sentences (1)-(4), must cohere with a speaker's intuitions about the truth values of those sentences.

The previous construal of the general methodological constraint for truth conditional theories of meaning is justified both by the facts concerning what any semanticist ought to take as data for their hypotheses, as well as the particular facts concerning the link between meanings and truth values on a truth conditional theory. First, the fact is that hypotheses about the meanings of sentences presuppose that those offering the hypotheses do not understand the language under study, but that those who are competent speakers of that language do understand it. This clearly rules out any direct linguistic communication between the two groups, which makes the overt linguistically-directed behaviors of fluent speakers one of the only guides for coming to know if a particular semantic hypothesis is correct. Second, on truth conditional theories, there is a conceptual link between meanings and truth values. On these theories, a sentence's truth value is determined by its meaning — its truth condition. A sentence gets the value true just in case the conditions for the truth of that sentence are satisfied, and gets the value false otherwise.

For the previous reasons, a truth conditionalist ought to infer that those fluent in the language — those who know the truth condition a given sentence has as its meaning — will be in a position to know which truth value to assign to a particular sentence, given a certain state of affairs. Assuming that speakers are more or less epistemically rational, and not generally deceitful, we can infer that they will assign true only to those sentences whose truth conditions are satisfied, and false to those that are not.

A speaker's truth value assignment to a sentence, then, for a truth conditionalist, ought to be a key piece of evidence in assessing semantic hypotheses. This is because it follows that, if a speaker is rational, and fluent in a particular language, then their assignments of truth value to sentences will reflect semantic content. Therefore, if a fluent speaker assigns true to a sentence, but some hypothesis predicts that it should be false, it is the hypothesis that is wrong, not the speaker. However, suppose the non-literalist insists on ignoring or dismissing the previous kind of evidence. Well, they then lose one of the primary means of testing their hypotheses. All that appears to be left are the requirements that hypotheses be mutually consistent and relatively simple — fairly weak requirements unless serving as a mere tie-breaker between hypotheses equally well supported for by other evidence. Giving up on natural truth value assignments as source of evidence for a truth conditional hypothesis, then, threatens to make truth conditional theories of meaning ad hoc, vacuous, or unfalsifiable.[11]

Non-literalism does not respect the previous methodological constraint, at least for fictive discourse. This theorist then must offer an explanation of some kind for why, in the case of fictive discourse, there is a justifiable exception to the methodological coherence

[11] See Stojanovic (2012) for the claim that such a principle is in fact required for truth conditional approaches to remain scientifically respectable.

constraint. In contrast, literalism clearly respects the constraint, thereby ensuring that truth conditional approaches do not become untethered from any data-driven considerations, and therefore maintaining their scientific respectability. What at first seems like a weakness for literalism — rejecting a strongly held semantic thesis — upon closer examination, turns out be a strength.

6. Objections and Replies

One potential non-literalist objection addresses my argument that it cannot cope with fan fiction and fictive modal discourse by offering non-literalist accounts of these two types of discourse. I consider possible analyses, but argue that they are still insufficient. A different, more general objection relies on arguing that it is a mistake to assume that the values assigned to sentences like (1) and (2) are in fact truth values, as opposed some other values, such as the values of being appropriate or inappropriate.

6.1 A Non-literalist Account of Fan Fiction

I will now return to re-examining fan fiction — originally presented as challenging the story operator view. As it turns out, there is an account of fan fiction that the non-literalist could offer, but it requires moving beyond a simple story operator account. It requires invoking another notion — a game of pretence (Walton, 2006). Games of pretence involve imagining that things are not exactly as they are, and sometimes feature props, such as works of fiction, whose properties constrain the game without wholly determining it. Pretence can explain fan fiction because it can be understood as constituted by a set of elaborate games of pretence that use works of fiction as props. In these cases, what is true in a particular game will be limited by what is true in the particular story that serves as that game's prop. However, a work of fiction serving as a prop in a game of pretence does not require that

everything true in that game is the very same as what is true in that game's related story. In fact, it is not even required that what is true in a particular story must be true in the game using that story as a prop. So how does what is true in a story constrain what is true in a game? We can think of the relation this way: what is true in a game of pretence is an extension of what is necessarily true in a story. Fan fiction simply exploits the looseness — the lack of an isomorphic relation — between what is true in a game and what is true in a story.

By invoking the notion of a game of pretence, this hybrid account (Devitt, 1981) can explain our sense that there is truth-apt fictive discourse that is not about what is true in a story without ever needing to relinquish the non-literalist stance. Those fictive statements drawn from fan fiction, which go beyond what is true in a story, can be explained in virtue of what is true in a game of pretence — still a non-literal context.

Nevertheless, this approach is subject to a specific kind of counterexample. Consider the fact that the character James Bond has been depicted in various ways, emphasizing certain traits and downplaying others in different stories and genres. Imagine now that two speakers are engaging in a debate about which depiction accurately represents the true or real Bond. This debate is not about what is true in the story, since in most of these cases, the various depictions are consistent with the story, or at least with what is necessarily true within it. More importantly, however, is that this debate is also not about what it is appropriate to pretend, since it is in fact appropriate, even required, to pretend that all of the various depictions depict the same character. There is reason, then, to doubt that any hybrid pretence-story operator account can successfully explain all forms of fictive discourse.

6.2 A Non-literalist Account of Fictive Modal Discourse

The second type of discourse that challenged the non-literalist's account — fictive modal discourse — also has a potential non-literalist explanation. This time, the explanation does not appeal to any notions not already accepted by the story operator theorist. According to this non-literalist, our intuition about modal discourse — that fictional characters have essential properties that are not essential to them according to the story — is explained in virtue of the scope interactions between story operators, and necessity or possibility operators.[12] On this proposal, the claim that Sherlock Holmes is essentially a detective is true because the sentence 'Necessarily, in the story, Sherlock Holmes is a detective' is true. In contrast, the opposing claim — that Sherlock Holmes is not essentially a detective — is true because the sentence 'In the story, necessarily, Sherlock Holmes is a detective' is false.

At least two issues arise with respect to this proposal. The first is how to determine the modal properties of a story. The second concerns whether the modal properties of stories align with the modal properties of fictional characters. Resolving the first issue involves finding a plausible account of a story's identity conditions. Resolving the second issue involves comparing the modal properties of stories with the modal properties of fictional characters. If the non-literalist does not offer an account of a story's modal properties, the project of reducing fictive modal claims to claims that rely only on the interactions between various operators will be incomplete. Worse, if the modal properties of stories fail to align with the modal properties of fictional characters, then the project simply fails. There can be no reduction of the modal properties of fictional characters in terms of the modal properties of stories.

---

[12] Thanks to Adam Sennet for this objection.

To begin exploring the first issue — that of the identity conditions on stories — consider the idea that a story is the same story in other possible scenarios just in case its content remains the same in those other scenarios. This idea, which I'll call the "content criterion" of story identity, would make it true that necessarily Sherlock Holmes is a detective. However, it would also make every other truth in the story necessary, which clearly seems too strong. Our ordinary practices, for instance, concerning fairytales, myths, or orally told stories, allow for some variations in content without a loss of story identity. The content criterion then is too narrow, failing to count a story as maintaining its integrity in cases in which, intuitively, it does.

Moreover, the condition is also too broad, counting stories as identical that are intuitively distinct. For example, it is possible to have two stories meet the content criterion, but which originate from separate independent sources. On the content criterion for story identity, this entails that these two sources have produced the very same story. However, various questions that arise reveal that this result has some potentially implausible consequences. For instance, imagine that a story is an entity created in a certain time and place, and suppose that two authors wrote one and the same story. Must they then have overlapped during its creation? Or, suppose one author bases a story, which is identical in content to another, on that author's very own experiences, but that the other story originates completely from the other author's imagination. If the stories now count as the same by the content criterion, does this entail that the story is both about events in the actual world and not? Addressing such questions is no easy task, and the non-literalist might do well to consider an alternative identity criterion.

One alternative is that a story's identity is tied to its origin — the origin criterion. That

is, a story's identity depends upon its having been created by a particular author at a particular time and in a particular place. This might allow for a story's content to vary in different scenarios without compromising its identity, and it would not count stories that have the same content, but that originate from different sources as the same story. So far, so good. But what modal properties does such an account entail are true of stories in general? Presumably, on the origin criterion, a story's modal properties will be derived from the modal properties of the event of its writing by its author. However, this would mean that this criterion is also too broad, or permissive. For instance, consider the fact that it appears possible that Emily Bronte, at the very same time and place at which she wrote *Wuthering Heights*, could have written a radically different story. Relying on the origin criterion, it turns out that because these two story counterparts are identical in their origins, they will count as one and the same story. But this is absurd. While it is true that a story's content may vary to a certain extent without compromising its identity, its content cannot vary radically without doing so. The specific constraints needed to get this result are not obvious.

Worse still, for the non-literalist's reductionist project, is that our intuitions concerning the modal properties of fictional characters do not in fact track the modal properties of stories. Note, for instance, some of the modal properties of the fictional character Heathcliff from *Wuthering Heights*. It seems entirely possible, for instance, that Heathcliff might not have tortured Isabella, owned several vicious canines, or died of a broken heart in front of an open window with the rain pouring in upon him. In contrast, it does not seem possible that he could have failed to have loved Catherine Earnshaw, or that he might have taken her perceived rejection of him in stride. Heathcliff, the character, just isn't built that way.

On the non-literalist's account, what would make the previous claims about Heathcliff

true would be the modal properties that hold of *Wuthering Heights*, which we can represent by giving the relevant necessity and possibility operators wide scope over the relevant story operators. Applying this account, if the previous modal claims are true of Heathcliff, it must be true that possibly, according to *Wuthering Heights*, Healthcliff never tortures Isabella, that he does not own any vicious canines, or that he does not die in front of a window. It must also be true that necessarily, according to *Wuthering Heights,* Healthcliff loves Catherine Earnshaw, and that he does not take her perceived rejection in stride. Are these facts true? I claim that the first three are not, which is sufficient to entail the failure of the non-literalist's reductionist project.

The argument — that a character's modal properties do not align with a story's — relies on the observation that a story's nature, meaning, and interpretation, is audience-directed in a way that the modal facts concerning fictional characters are not. That is, a story is intended to convey something to an audience, whereas characters are not. Story-telling is a kind of speech act. Fictional characters are not. As I will argue, this entails that the modal properties of fictional characters do not correspond to the modal properties of stories.

Consider, for instance, the plot-driving elements of a story — its characters, or their roles at least, and their interactions — and whether they are essential to it, and likewise for its atmosphere, and thematic elements. Imagine, for example, that Isabella or an Isabella substitute, had never existed, and that therefore Heathcliff's level of depravity had not been revealed in the story by his actions towards her. Or consider the idea that the dogs at Wuthering Heights were actually friendly, that the weather was always fair, and so on. It seems that changing these particular facts about *Wuthering Heights* would make for a very

different kind of story, one that was not concerned with the constant threat, and eventual triumph, of our unbridled, uncivilized, and more natural passions over our rule-oriented, civilized, and less natural desire for conformity. Arguably, this message is an essential feature of the story *Wuthering Heights.* In contrast, it is not plausible to suppose that a fictional character's modal properties, their fictive modal properties at least, have much, if anything, to do with the audience directed elements of a story. These facts show that the very idea of giving a non-literalist reductionist analysis of fictive modal discourse is deeply mistaken.

6.3 Judgements of Literal Truth vs. Judgements of Appropriateness

A final way to reject literalism is to reject the claim that speakers are using the concept of truth at all in assigning values to sentences like (1) and (2). Instead, they are using the different concept of what it is appropriate to say. For instance, it is appropriate to say that you are fine when asked about your well-being, regardless of the actual status of your well-being. The concept of truth does not guide the appropriate response in this case. According to this line of reasoning, all forms of discourse have rules of appropriateness. Fictive discourse should be no exception, and the presence of these rules should be no cause for alarm. The objection is that the literalist mistakenly infers that because there are rules of appropriateness governing fictive discourse, it is therefore truth-apt. If this is correct, if fictive discourse is not in fact truth-apt, then the puzzle of fictive discourse is illusory.

To respond to the previous line of reasoning, the literalist can agree that, even if all forms of discourse are governed by rules of appropriateness, which do not inherently appeal to the concept of truth, there are nevertheless certain kinds of saying that do appeal to norms of truth. Take, for example, acts of assertion. Plausibly, the felicity conditions on

acts of assertion do involve the concept of truth, constitutively. But, the example that was offered, having to do with an appropriate response to a certain kind of query, was not an example of what it is appropriate to assert, only of what it is appropriate to say. So, of course, its felicity conditions do not necessarily involve the concept of truth. What it is appropriate to say can be governed by all sorts of different norms. In the example considered, the reason it is appropriate to say that you are fine is that certain social conventions dictate that you should not burden others with your problems by mentioning them, especially to mere acquaintances. But this tells us nothing about what it would be appropriate to assert. In sum, the rules that govern the appropriateness of saying are not the same as the rules that govern the appropriateness of assertion. And, in fact, there are good reasons for taking much of fictive discourse as an instance of assertive discourse. For instance, whether it is appropriate to say that Sherlock Holmes smokes will depend upon whether what is said accurately reflects what is contained within a work of fiction. If it is accurate, then what was said is true, and if it is not accurate, then what was said is false. Standards of representational accuracy, then, invoke the concept of truth. Since standards of representational accuracy govern much of fictive discourse, much of it will count as assertive. Therefore, the assignment of values to sentences like (1) do, in fact, rely on the concepts of truth and falsity. It follows that the puzzle of fictive discourse is not illusory after all.

7. Concluding Remarks

I have argued that our assignment of the values to sentences like (1)-(4), as well as others, is not fully explained by the non-literalist accounts considered. Neither is the fact that such assignments are robust, persisting even in neutral contexts. Furthermore, non-literalism

also cannot explain the role that truth value assignments play, in general, in a scientifically respectable truth conditional theory of meaning. For all of these reasons, I claim that literalism is the most reasonable approach for analyzing fictive discourse. That is, if we are truly interested in what is said by speakers when engaged in fictive discourse, we ought to take our truth value assignment to a sentence like (1) at face value, as literally true, and therefore as indicative of the semantic content of those sentences.[13]

SUNY Geneseo

Bibliography

Austen, Jane (1993). *Pride and Prejudice*. Hertfordshire: Wordsworth Editions Limited.

Bertolet, Rod. (1984). "On Fictional Ellipsis." Erkenntnis 21(2): 189-194.

Braun, David. (2003). "Empty Names, Fictional Names, Mythical Names". Noûs 39(4): 596–631.

Bronte, Emiy. (1976). *Wuthering Heights.* Oxford: Oxford at the Clarendon Press.

Carnap, R., 1947. *Meaning and Necessity*. Chicago: The University of Chicago Press.

Currie, Gregory. (1988). "Fictional Names." Australasian Journal of Philosophy 66: 471-488.

Davidson, Donald. (1967). "Truth and Meaning." *Synthese*,17: 304–23.

Deutsch, Harry. (2000). "Making Up Stories." (In Anthony Everett and Thomas Hofweber (Eds.), Empty Names, Fiction and the Puzzles of Non-Existence:149-181. Stanford, CA: CSLI.

Devitt, Michael. (1981). Designation. (New York, NY: Columbia University Press)

Doyle, Sir Arthur Conan. (2002). The Complete Sherlock Holmes. New York: Gramercy Books.

Elgin, Catherine. (1983). With Reference to Reference. Indianapolis: Hackett Publishing Company.

---

[13] Thanks to John Horty, Peter Ludlow, and Adam Sennet for comments on earlier drafts. Thanks also go to anonymous journal referees.

Evans, Gareth. (1982). The Varieties of Reference. Ed. John McDowell, Oxford: Oxford University Press.

Frege, Gottlob. (1956). “The Thought." *Mind* 65/259: 289-311.

Grice, Paul. (1975). "Logic and Conversation". In Cole, P. and Morgan, J. (eds.) Syntax and semantics, vol 3. New York: Academic Press.

Max Kolbel, (2001). “Two Dogmas of Davidsonian Semantics.” Journal of Philosophy. xcviii, 12: pp. 613-635.

Kripke, Saul A. (1980). Naming and Necessity. Cambridge, MA: Harvard University Press.
--------.(1973, ms). Reference and Existence: The John Locke Lectures for 1973.

Lewis, David. (1986). *On the Plurality of Worlds*. Blackwell: Oxford.
————. (1978). “Truth in Fiction.” American Philosophical Quarterly 15.1: 37-46.

Ludlow, Peter. (2006). “From Sherlock and Buffy to Klingon and Norrathian Platinum Pieces: Pretense, Contextualism, and the Myth of Fiction.” Philosophical Issues 16.1: 162-183

Martin, Robert M. and Peter Schotch. (1974). “The Meaning of Fictional Names.” Philosophical Studies 26: 377-388.

Piccinini, Gualtiero and Sam Scott. (2010). “Recovering What is Said with Empty Names,” Canadian Journal of Philosophy 40(2): 239-274.

Parsons, Terence. (1980). Nonexistent Objects. Yale University Press.

Recanati, Francois. (2004). Literal Meaning. Cambridge University Press.

Russell, Bertrand. (1905). “On Denoting.” *Mind* 14 (56): 479–493.

Sainsbury, Mark. (2009). Fiction and Fictionalism (New Problems of Philosophy). Routledge.

Stojanovic, Isadora. (2012). “Domain-sensitivity.” Synthese (Relativity) 184(2): 137-155.

Szabo, Zoltan, Gendler, Ed. (2005). Semantics versus Pragmatics. Clarendon Press.

Thomasson, Amie. L. (2008). Fiction and Metaphysics. Cambridge Studies in Philosophy.

Tiedke, Heidi. (2011). “Proper Names and Their Fictional Uses.” Australasian Journal of Philosophy 89(4): 707-726.

Travis, Charles. (1996) “Meaning’s Role in Truth” Mind 105: 451-466.

VanInwagen, Peter. (1977). “Creatures of Fiction.” American Philosophical Quarterly 14/4: 299-308.

Walton, Kendall. (2006). Mimesis as Make-believe: On the Foundations of the Representational Arts. Harvard University Press.

Zalta, E. (1983). Abstract Objects: An Introduction to Axiomatic Metaphysics. D. Reidel