

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

Heidi Savage
SUNY Geneseo

Abstract: Many, if not most philosophers, deny that a sentence like ‘Sherlock Holmes smokes’ is true. However, this attitude conflicts with speakers’ assignment of the value true to this sentence. Furthermore, making these assignments seem in no way distinct from the process that leads speakers to assign true to other sentences, sentences like ‘Bertrand Russell smokes.’ I will explore the idea that when speakers assign the value true to the first sentence, they are not making any kind of confused mistake — that we ought to take these assignments at face value. I show how the alternative view is inadequate for explaining various examples of fictional discourse. In addition, evidence that these truth value assignments to sentences are tracking semantic content, rather than pragmatic effects, is offered.

1. Introduction

Speakers often engage in discourse about characters from works of fiction, which seems both meaningful and able to convey truths. This occurs not only within the context of theoretical analysis and criticism, involving abstract literary properties, but also within more ordinary contexts, involving simple everyday properties. The latter I call “fictive discourse” to distinguish it from the former kind, commonly known as “meta-fictive discourse.” Instances of sentences from fictive discourse include those such as ‘Sherlock Holmes smokes’, ‘Heathcliff is a tyrant’, and ‘Elizabeth Bennet has a sharp wit’ — drawn respectively from the fictional works written by Doyle (2002), Bronte (1976), and Austen (1993).

Now each of these sentences appear to express something true. However, speakers will also commonly endorse the idea that if a sentence is true that is because something it accurately represents something about the world that makes it true (Moltmann, 2018). For example, in understanding that the sentence ‘snow is white’ is true, speakers assume that there must be some substance snow and that it has the property of being white. However, most speakers would also reject the idea that the former fictive sentences are, respectively, about the individual pipe-smoking Sherlock Holmes, tyrannical Heathcliff, or the witty Elizabeth Bennet. For instance, speakers might utter the following sentence ‘Sherlock

Holmes was a pipe-smoker, which is a bad habit, but he was just a fictional character, and doesn't exist anyway.' In other words, speakers know that the previous sentences, instances of fictive discourse, are not about anything, and yet their intuition — that they are true — remains.¹

The previous discourse clearly raises a puzzle in need of resolution, which I will call the "puzzle of fictive discourse." It poses a challenge for almost any widely accepted theory of meaning: Fregeanism, Russellianism, Davidsonian, Possible Worlds Semantics, and their variants. All these theories count, broadly speaking, as truth conditional in nature, characterized by the belief that the meanings of declarative sentences consist in their informative, or truth-apt content.

Furthermore, encoding speakers' common assumptions, these theories also typically adopt the evaluative semantic principle that simple subject-predicate sentences, specifically those containing names, are true just in case the referent of the name has the property associated with that sentence's predicate. This is what I will call the "traditional" analysis of simple subject-predicate sentences containing proper names. It predicts that all fictive sentences are false, which conflicts with the intuition that they are true. Since on this analysis, if there is no Sherlock Holmes, Heathcliff, or Elizabeth Bennet to serve as the referents for the corresponding names, then any sentence that purports to be about any of these individuals, cannot truly be about them at all, since they do not exist.

Truth conditionalists commonly explain the intuition that certain fictive sentences are true in virtue of being associated with other kind of content. This is one way to deal with the puzzle of fictive discourse – reject the intuition that the previous sentences express anything

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

true at all. This is what I call a “non-literalist” approach to the puzzle of fictive discourse.² In contrast, what I call a “literalist” approach to the puzzle, the approach I support, takes it seriously that the previous sentences do express something true in virtue of their semantic content.³ Adopting a literalist position, then, conflicts with the traditional analysis — a well-accepted, well-warranted account.⁴

If there is already an explanation available that is consistent with the traditional analysis, why not adopt it? 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 fictive sentences from fictive discourse — the story operator account — does not apply to certain types of fictive sentences. Some types of fictive discourse are not plausibly analyzed using a story operator, thereby making literalism an approach worth considering. The second argument offers positive support for adopting literalism since assignments of the value true to certain fictive sentences is steadfast even in a context that demands a literal interpretation of those sentences. The third argument shows that literalism, despite first impressions, respects truth-conditional theories of meaning better than non-literalism does.

In the next section, I describe the puzzle of fictive discourse in more explicit terms, as well as elaborating on the motivations for both non-literalist and literalist approaches to it. In section three, I offer negative reasons for accepting literalism -- that non-literalism cannot account for two distinct types of fictive discourse. Section four offers two positive reasons to adopt literalism that come from methodological considerations and because literalism passes a specific test I offer for diagnosing the literal content of a sentence. I then explain, in section

² see Currie (1988); Evans (1982); Lewis (1983); and Walton (2006), among many others.

³ see Deutsch (2001); Elgin (1986); Ludlow (2006); and Martin and Schotch (1974); Tiedke (2011).

⁴ I defend a specific literalist analysis of the content fictive discourse in previous work in which I reject the traditional analysis of its content in my Tiedke (2011).

five, how the non-literalist might respond to the critiques offered in section three, concluding that these responses are unconvincing.

2. The Puzzle of Fictive Discourse

To illustrate the nature of the puzzle of fictive discourse more clearly, consider the following 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, which are respectively, first true and then false, for each pair.⁵ That speakers do, in fact, have these intuitions gets support from a study conducted by Piccinini and Scott (2010) reporting that over 80% of informed speakers will assign the value true to sentences like (1). Presumably, they would not assign that value to sentence (2). Likewise, we should expect that informed speakers will repeat this pattern of assignments for sentences (3) and (4). Furthermore, these assignments to sentences like (1) and (2) occur as spontaneously and as effortlessly as assignments for sentences like (3) and (4). For Piccinini and Scott, this fact suggests that the procedure for assigning truth values is the same in both cases, and that therefore, our treatment of both pairs of sentences should be uniform.⁶

⁵ There is a third position on the truth value of sentences like (1) and (2). This position denies that they have any truth value at all. They are instead indeterminate. While I am not opposed to this idea, at least not on the 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).

⁶ 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).

Regarding 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 the types of basic syntactic parts constituting a particular sentence, the types of meanings associated with those parts, and the rules for combining those meaningful syntactic parts to form more complex types of meaningful phrases. The traditional analysis of the content of a simple subject-predicate sentence is a good example of a compositional semantic rule. Now, since sentences (1)-(4) do not apparently differ with respect to the basic types of syntactic parts they contain, or their ways of being combined, they should all get the same semantic analysis, and this is one of the underlying assumptions that generates the puzzle of fictive discourse.

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 — that four plausible claims are mutually inconsistent, and that therefore, one must be given up.

I will not explore the idea that we can give up compositionality, or that sentences (1) and (2) might have referents. The reason for refusing to give up compositionality is simple. It threatens the entire project of giving any systematic analysis of natural language. The reason for refusing to posit referents 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 fictive sentences like (1)-(4).⁷

⁷ Many philosophers, including Kripke (1973 ms.), VanLwagen (1977), and Zalta (1983),

The only option, then, is to either deny that the assignments of values to sentences (1) and (2) reflect evaluations of the literal semantic content of those sentences as non-literalists do, or to reject the traditional analysis of that content as literalists do. Defending either of these positions is difficult.

Because the non-literalist accepts the traditional analysis, they must explain why speakers assign the value true to a sentence like (1), as well as why sentences like (1) and (2) would even get different truth value assignments at all, since their view entails that both are equally false.⁸ Literalists face a different challenge – that 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.

Having clarified the puzzle of fictive discourse, and outlined the two main positions with respect to it, I will now describe and illustrate the failure of the non-literalist's typical response to the puzzle – the story operator theory. This motivates looking for an alternative approach. In fact, the way in which the story operator theory fails suggests not just another approach, but specifically a literalist approach.

3. Problems for The Non-literalist Story Operator Account

The standard non-literalist solution to the puzzle of fictive discourse – the story operator account (Lewis, 1978) – holds that sentences like (1) and (2) appear have truth values only because they are pragmatically associated with other sentences that are true. which are

among others, posit abstract objects — fictional characters – as referents of fictional names. This approach is motivated by considering a different type of fictional discourse – meta-fictive discourse – that also contains sentences that appear to be true such as 'There are more characters in *Pride and Prejudice* than in *Wuthering Heights*.' For arguments that positing referents does not help even with that problem, see Garcia-Carpintero (2019) and Savage (2020).

⁸ For explanations of this fact, see Bertolet (1974), Braun (2003), Evans (1982). Lewis (1978), and Walton (2006).

versions of the previous sentences, except that they are qualified with a story operator like 'In the story' or 'According to the story.' On this account, then, sentence (1) appears true because sentence

(1)' In the story, Sherlock Holmes smokes

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

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

is false. The fact that speakers assign different truth values to sentences (1) and (2) is explained in virtue of the difference between the truth values of other sentences that they are associated with – those qualified with a story operator.

While the semantic content of sentences (1)-(4), for the non-literalist, is given by the traditional analysis, and therefore (1) and (2) are not true, fictive sentences can be used to express certain true or false statements, because these sentences are not taken at face value, but are understood as qualified with a story operator. There are, at least, two distinct explanations for accepting that fictive sentences are not or should not be taken at face value. The first is that sentences (1) and (2) are ambiguous between qualified and non-qualified interpretations. This explanation, however, is inconsistent with non-literalism about fictive sentences since it entails that sentences (1) and (2) can be literally true or false. The non-literalist must appeal, then, to a second explanation for the claim that we should not take fictive sentences at face value. This second explanation is that it is because fictive sentences are pragmatically associated with sentences qualified with a story operator that some of them seem true (Bertolet, 1984).⁹

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

⁹ There are inherent problems with the first explanation at any rate (Savage, 2020).

the account its due. I then discuss two other conversations that reveal its shortcomings, 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. We have an initial literal traditional 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 case supports 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 of fictive discourse. Let us assume that Sam again believes that it is impossible that a person, fictional or otherwise, could use mind altering substances to such the extent that Holmes did and be as successful as he was.

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 don't think that Holmes could have done that and been such a great detective. Makes me wonder if maybe 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 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 fictive discourse that contains sentences whose truth values are clearly not determined by what is true or false according to a story.

The story operator account does not exhaust all forms of truth-apt fictive discourse, at least not fan fiction. However, fan fiction is not the only form of truth-apt fictive discourse that 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.

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

Fictive modal discourse is about the essential or non-essential properties of fictional characters. For example, the sentence ‘It is possible that Sherlock Holmes was not a detective’ is a sentence that counts as belonging to fictive modal discourse. As it turns out, claims from this kind of discourse can be interpreted in two different, but equally legitimate ways — in some literalist way inconsistent with the traditional analysis, and in the qualified story operator way.

To begin with, consider this hypothetical conversation, once again between Adam and Sam, concerning whether Sherlock Holmes could have been other than 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 some sense in which Holmes would not be Holmes if he was not a detective? I mean, what properties make Holmes what he is, outside of those decided upon by Doyle? Without the Doyle's act of penning the stories, Holmes would be nothing at all. And one the properties he deemed Holmes had was being a detective.

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 right way to answer this question — that Sherlock Holmes could not have been other than a detective. The problem that this modal intuition presents for the story operator account is that it seems that it could never come out true, on that view, 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.

But is this an intuition worth worrying about? It is if enough speakers share it. If enough speakers agree that sentence

(5) Sherlock Holmes could not have been other than a detective

is true, and if there is an explanation of its truth, then I would say it is worth thinking about. I will now suggest such an explanation, beginning with the explanation that starts in Conversation 3.

Adam's explanation, in conversation 3, appeals to the idea that fictional characters are only as real as the acts of their authors. For Adam, Holmes is essentially a detective is because Doyle created Sherlock Holmes via by acts of that stipulatively associating the name 'Sherlock Holmes' with certain properties. Since the name 'Sherlock Holmes' is stipulatively associated with being a detective, that property is then part of what makes up Sherlock

Holmes, and therefore Sherlock Holmes must essentially have the property of being a detective.¹⁰ At least one problem with this idea is that it applies to any property the author stipulatively associates with 'Sherlock Holmes', say for instance, the specific color of the socks he wears. So even though it is reasonable to think of Sherlock Holmes as having the property of being a detective essentially, it is not reasonable to believe this of the specific color of socks he wears.¹¹

Whatever kind of analysis is imposed on these stipulative acts involved the making of a work fiction – an ontologically heavy metaphysical interpretation that makes them acts of creating an entity called a "fictional character", or a more deflated ontological linguistic interpretation that makes them acts of creating or introducing a name with cognitive significance – I need to offer some explanation of the idea that Sherlock Holmes has being a detective an essential property, but not wearing black socks, let us say. I give my explanation relying on realist terms, but to show my faux commitment to realism, I use scare quotes quite heavily, and I assume annoyingly. I do this, however, for economic reasons, since this is only a small part of what the entire paper is about, and putting my explanation in ontologically deflated linguistic terms would take much more description, though this can be done, it will not be done so here.

The explanation I offer distinguishes between what makes something count as the

¹⁰ The metaphysical and linguistic aspects of creating a work of fiction, of maintaining these distinctions, of heeding Quine's lessons (1936) about the role that stipulation or convention might play in an explanation of fictive discourse, are all important issues in need of close analysis and explanation. I discuss these ideas briefly here, and attempt a careful usage of them, but I do not offer a developed account of the nature of the role of names, stipulations, and metaphysics, in the creative activity of fiction writing in what follows. The ideas discussed I put forward tentatively with the hope of further development in later work.

¹¹ First, notice that the explanation itself assumes that if you are created with a particular property that this must be part of your essence, which is not obvious, as I will next discuss. Second, note also that there may be properties readers take to be essential of a fictional character that were never directly stipulatively associated with its moniker.

kind "fictional character" from what makes said fictional character that specific character. That is, a "fictional character" might exist simply by being mentioned by name in a work of fiction, but this tells us nothing of that character's character – its essential properties as a unique "fictional character." For instance, what makes Sherlock Holmes a "fictional character" is the fact that an author "created" him via some act(s) of stipulatively associating certain properties with the name 'Sherlock Holmes.' While all these acts of stipulation are part of what "create" the so-called fictional character Sherlock Holmes, not all of them have to form part of Sherlock Holmes's "essential properties," only a certain subset of them do. This subset of stipulative actions will be those that determine Sherlock Holmes' character or nature — his essential properties.

I suggest that the stipulations used to determine the essential properties of a are those that the author chooses that fix that character's role in 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, but it is not essential that he wear black socks to play this role, and therefore that is not one of his essential properties.

I offered a sketch of an explanation for the truth-aptness of fictive modal discourse – another example of fictive discourse that the story operator model cannot handle in any obvious way. For instance, consider the fact that when I have asked speakers to evaluate the truth of a sentence like (5), they sometimes express confusion or uncertainty about the answer. But this confusion and uncertainty cannot concern whether to evaluate sentence (5) by relying on a story operator account or the traditional analysis — the two non-literalist options – since on either of these readings the sentence is clearly false.

In fact, fictive modal discourse – independently of the story operator analysis – shows that it poses a direct challenge to the traditional analysis. Why? Because when asked the

question of whether sentence (5) is true, speakers are never confused about whether to answer it relying on the traditional analysis. It is always an uncertainty about whether to evaluate the sentence within the story or outside of the story. Once stipulated that it is not what is true within the story that is at issue, speakers will express various opinions about whether sentence (5) is true. If the traditional analysis were correct, all speakers should agree that the sentence is clearly false, but they do not. The false traditional reading — what I think of as the “philosophers’ reading” — does not 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 at all in analyzing its content .¹² I offered an alternative suggestion about how to evaluate sentences like (5), but I did not offer an explicit evaluative rule for how to evaluate these sentences.¹³

Fictive modal discourse, then, presents both an 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.

4. Reasons for Being a Literalist About Fictive Discourse

While as admirable as the simplicity and plausibility of the story operator account might be, as we have seen, it cannot explain certain cases of this kind of discourse. This offers a negative reason for considering literalism. I will now give two positive arguments in its favor. The first rests upon methodological considerations. The second depends upon a test that

¹² 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.

¹³ I do offer a literalist rule (2011) for simple predicative sentences such as sentences (1) and (2), which I would need to be adapt for sentences like sentence (5).

supports literalist interpretations of sentences like (1) and (2).

4.1 The Methodology of Giving Semantic Theories

While non-literalists may face challenges, nevertheless, they support a commitment to the traditional analysis, which is plausible, entrenched, well-accepted, and well-warranted. However, I argue that, certain plausible methodological constraints on theories of meaning, which include truth conditional theories of meaning, proves that truth conditionalists should reject the traditional analysis of (1)-(4), thereby undermining the motivation for adopting a non-literalist approach to fictive discourse in the first place.

The general methodological constraints are based on the idea that any semantic theory of natural language ought to fit with speakers' linguistic behaviors, and that if they fail to do so, we ought to reject it, or to engage in its revision. These constraints are two-fold: (a) the theorist giving a theory of the meaning of a particular natural language either does not know the meaning of that language, or must screen off that knowledge in theorizing about the language that they choose to study; (b) Assuming the native speakers of the language under study do in fact understand or know this language, their overt linguistic actions ought to serve as one of the main guides for testing a theorist's semantic hypothesis.¹⁴ In sum, (a) and (b) entail a coherence requirement on any theory of meaning. That is, theoretical hypotheses and speakers' overt linguistic behavior need to "mesh."¹⁵

Concerning truth conditional theories of meaning, a conceptual link exists 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

¹⁴ Though what knowledge speakers in fact is controversial according to Stich (1971).

¹⁵ Denying that this is the case leaves semantic hypotheses about a language free-floating from any data, since I see now other stronger more plausible source of such data than from natural language speakers themselves. They risk therefore becoming ad-hoc or unfalsifiable.

the truth of that sentence are satisfied and gets the value false otherwise. Methodological constraints (a) and (b) would require giving a significant role to the intuitions of native speakers concerning the truth values assigned to sentences of the language under study.

Given (a) and (b), if competent speakers assign the value true to certain sentences, this must be because their understanding of its meaning – its truth condition – must be satisfied. That is, 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 particular 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. If a fluent speaker assigns true to a sentence, but the truth conditionalist hypothesis predicts that it should be false, this is evidence for the falsity of the hypothesis.

The non-literalist in their refusal to take the assignments of the truth values that speakers in fact give to sentence (1) and (2), and in remaining committed to the traditional analysis, violate the methodological constraints (a) and (b). In doing so, they lose one of the primary means of evaluating their hypotheses. All that appears to be left to constrain truth conditional theories are the requirements of mutual consistency and of simplicity — relatively weak requirements.

Giving up on natural truth value assignments as source of evidence for a truth conditional hypothesis threatens to make non-literalist truth conditional theories of meaning ad hoc, vacuous, or unfalsifiable.¹⁶ In contrast, literalism clearly respects the constraint,

¹⁶ See Stojanovic (2012) for the claim that such a principle is in fact required for truth conditional approaches to remain scientifically respectable.

thereby ensuring that truth conditional approaches do not become untethered from any data-driven considerations, and therefore maintaining their scientific respectability. We should therefore reject the traditional analysis, which entails the wrong predications about speakers' overt linguistic actions, and instead develop a literalist analysis of fictive discourse.

4.2 The Neutral Context Test

The non-literalist will point out that constraints (a) and (b) are not absolutes. Sometimes linguistic phenomena are a function of the pragmatics of communication, rather than literal content, as the non-literalist's appeal to story operator account is intended to be applied. It explains our intuitions about the truth values of sentences (1) and (2) as the result of not vocalizing the presence of a story operator in uttering these sentences for reasons of convenience of expediency.

However, there are tests that are supposed to isolate a sentence's semantic content from any pragmatic content with which it might be associated. Such tests try to avoid embedding a sentence within a conversation, so that the triggering of pragmatic effects is unlikely. If a test sentence occurs in a context that does not trigger any pragmatic effects, then any interpretation of that sentence must convey its literal content.

One test that does not trigger pragmatic effects I call the "neutral context" test.¹⁷ It is performed by asking speakers to assign a truth value to a test sentence that someone merely mentions, rather than uses. This might also occur, for instance, in a context in which an instructor displays a sentence on a chalkboard, and speakers then evaluate it for truth, as I have done many times using sentences (1) and (2). In these contexts, we can safely assume that a speaker's assignments of truth values would be tracking semantic, rather than

¹⁷ There are several tests used to do so. One is the courtroom test; another is the betting test. Neither is conclusive.

pragmatic content, and theorists would have to respect constraints (a) and (b).

If displaying sentences on a chalkboard counts as an instance of a neutral context, then such displays of sentences (1) and (2), which nearly universally garner the assignments true then false respectively, this shows that the neutral context test favors literalism about fictive discourse. In fact, I claim that the original presentation of sentences (1)-(4) earlier is an instance of presenting sentences in a paradigm neutral context, and the natural truth value assignments, once again, provide evidence in favor of a literalist approach to the puzzle of fictive discourse. These results then offer another positive reason for adopting a literalist position.

Questions about the neutral context test's scientific validity may arise, however. Does mentioning a sentence truly isolate it from triggering pragmatic effects? To show that the test is scientifically valid, I must find a sentence with well-known literal content that in non-neutral contexts is rarely the basis of its assigned a truth value – a sentence that has an entrenched pragmatic interpretation – but that that content is the basis of assigning its truth value in a neutral context.

At least one sentence with an entrenched pragmatic interpretation is the sentence 'The patient is not going to die.' Rarely do speakers and hearers ever accuse a medical practitioner of having said something false because the patient is in fact mortal, contrary to the sentence's literal content. Instead, these utterances are evaluated as true or false on the grounds of whether the patient is going to die within some specified timeframe.

But now consider whether this everyday interpretation will survive the neutral context test. Consider what truth value speakers might assign the sentence 'The patient is not going to die' if simply displayed in a neutral context, outside of any conversational context. I predict that the pragmatic interpretation would not survive – that the sentence would get the value false on the grounds of that its literal interpretation is false.

If the neutral context test can induce a literal interpretation even in cases like the one considered — in which the sentence is rarely interpreted literally — it has proven its strength as an instrument for detecting differences in semantic and pragmatic content. The test's scientific validity is sound, and its results justify a literalist approach to fictive discourse.

5. Objections and Replies

I will now consider a non-literalist response to my argument that the story operator account cannot cope with fan fiction and fictive modal discourse. The response is simple: it can, with some modifications.¹⁸ I also consider a more general response to the argument that fictive discourse poses any kind of puzzle at all. It does so only if we 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. I offer arguments against both lines of response.

5.1 A Non-literalist Account of Fan Fiction

While the non-literalist can offer some kind of account of fan fiction, it requires moving beyond a simple story operator account and invoking another notion — a game of pretense.¹⁹ Games of pretense 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. Pretense might explain fan fiction, since it could be constituted by a set of elaborate games of pretense that rely on works of fiction as props. What is true according to the story, in these games, plays the same role that real-world facts would play in a game of mud-pie. So while it is true that according to the stories, Sherlock Holmes was a great detective, fan fiction writers take this "fact" into consideration in light of his drug habit, and

¹⁸ Thanks to Adam Sennet for the suggested modifications.

¹⁹ For developments of this view, see Evans (1982), and Walton (2006).

imagine that because of that habit, it is impossible that he could have performed the feats that are true according to the stories. While any piece of fan fiction, understood as a game of pretense, will be constrained by its props – works of fiction – specifically by what is true according to the story contained within those props, the constraints will allow for some "poetic license" for a fan fiction writer – not everything true in a story must be true in a game of fan fiction. So, then, how does what is true in a story constrain what is true in a game of fan fiction? We can think of the relation this way: what must be true in a game of pretense is an extension of the essential properties of a given stories. This might include atmosphere, characters, certain events. It would depend on the work of fiction in question.

By invoking the notion of a game of pretense to account for fan fiction, a hybrid account suggested by Devitt (1981), there is an explanation for 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. The fictive discourse that composes fan fiction, which goes beyond, and sometimes conflicts with what is true in a story, can be explained in virtue of what is true in a game of pretense, which is still a non-literal context. However, I offer a counterexample to this approach showing that not even a hybrid account can exhaust fictive discourse, fan fiction or otherwise.

Consider the various depictions of the character James Bond in different stories, some of which emphasize certain traits and downplay others – a bit like what the writers of fan fiction might do. Nevertheless, readers assume that there is a single character James Bond. Imagine that these different depictions lead two readers to engage in a debate about which one of them accurately represents the real James Bond. Certainly, this debate is not about what is true in any Bond-story since the fact that these traits are true of him according to the various stories is what initially generates the debate. Maybe, then, it is a debate that occurs within the realm fan fiction. But if the hybrid account of fan fiction is correct – that we should

analyze it in terms of games of pretense constrained by what is true according to certain stories – then there is no substantive debate to be found here, since it is of course proper to pretend that all of the various depictions depict the same character, even if they conflict. Assuming the hybrid account, this is all the fan fiction writer could say. But this was not the question in which the readers had taken an interest. This example, then, shows that the hybrid account of fan fiction either fails to account for all fan fiction, or that there is yet another type of fictive discourse that is in need of further analysis. Either way, the non-literalist is once again in trouble.

5.2 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 for which the needed result would be that fictional characters cannot have modal properties without being preceded by a story operator – a reductive project. To attempt this task, the non-literalist need not appeal to any notions not already accepted by the story operator theorist. According to this non-literalist, some of our intuitions about modal discourse — that fictional characters have modal properties that they do not have in the story — is explained in virtue of the scope interactions between story operators, and necessity or possibility operators.²⁰ On this proposal, the intuition that sentence (5) is true is explained by the fact that it is necessary in the Sherlock Holmes stories that Holmes could not have been other than a detective. In contrast, the intuition that sentence (5) can have a false reading is explained by the fact that it is false that in the story Sherlock Holmes could not possibly be something other than a detective.

At least two issues arise with respect to this proposal. First, in explaining our intuition about the truth of sentence (5), the necessity operator must take scope over the entire story,

²⁰ Thanks to Adam Sennet for this objection.

making such claims about the modal properties of stories, not characters within stories – the hoped for result on the part of the non-literalist story operator theorist. This requires finding a plausible account of a story's identity conditions, however. Assuming the first issue can be resolved, the account must then also make the correct intuitive predictions concerning the modal properties of a story's inhabitants -- "fictional characters."

If the non-literalist does not offer an account of a story's modal properties, the project of reducing modal claims about fictional characters in favor of 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.

There are at least two ways to understand the identity conditions on stories. First, consider the idea that a story is the same story in different possible scenarios just in case its content remains the same in those other scenarios. This idea, which I 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 insignificant variations in content without a loss of story identity. The content criterion then is too narrow. For instance, a single story about Sherlock Holmes would count as two distinct stories if in within the single story it says that he always wears black socks, and if presented in a play he always wears grey socks.

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 somehow written a single story. Besides its

sounding odd to start with, this consequence also has further implausible consequences. For instance, imagine that a story is an object of some kind that is created in a certain time and place, and suppose it makes sense that two authors wrote stories content-identical, and therefore wrote a single story, one and the same story, then it seems that the two authors must have overlapped in space-time 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 author's story originates completely from their imagination. If the stories now count as the same by the content criterion, this would entail that this single story is both about events in the actual world, and not about events in the actual world. Whether such issues could even be addressed is doubtful. The non-literalist would therefore do well to consider an alternative criterion on story identity.

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. If correct, however, 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 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 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, is not built that way.

On the non-literalist's account, what would make the previous claims about Heathcliff true would have to be the modal properties of the story *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, *Wuthering Heights* is a story in which Heathcliff 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*, Heathcliff 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. I will, however, also show that stories have modal properties that fictional characters do not.

The explanation the misalignment of a fictional character's modal properties with a story's is that a work of fiction's nature, meaning, and interpretation, are 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.

Consider, for instance, the symbolic elements of a story, and consider whether they are essential to it -- its atmosphere, its thematic elements, and so on. 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. This message is an essential feature of the story *Wuthering Heights*. I conclude that these differences between the modal properties of fictional characters and of stories show that the very idea of giving a non-literalist reductionist analysis of modal discourse about fictional characters is deeply mistaken.

5.3 Judgements of Literal Truth vs. Judgements of Appropriateness

A final way to reject literalism is to reject the claim that speakers in making their assignments to sentences (1), (2), and now (5), are relying on the concept of truth at all. Instead, they are using a different concept – that of what it is appropriate to say. For instance, it is correct to say that you are fine when asked by an acquaintance 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 simply an illusion.

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 inaccurate, 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), (2), and (5) do, in fact, rely on the concepts of truth and falsity. It follows that the puzzle of fictive discourse is not illusory after all.

6. Concluding Remarks

I have argued that our assignments of truth values to sentences like (1)-(4), as well as others, are 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, and literalism can. For all of these reasons, I claim that literalism is a strongly viable 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 assignments to fictive discourse at face value, as literally true, and therefore as indicative of the semantic content of those sentences.²¹

Works Cited

- 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, Emily. (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.
- Evans, Gareth. (1982). *The Varieties of Reference*. Ed. John McDowell, Oxford: Oxford University Press.
- Frege, Gottlob. (1956). "The Thought." *Mind* 65/259: 289-311.

²¹ Thanks to John Horty, Peter Ludlow, and Adam Sennet for comments on earlier drafts. Thanks also go to anonymous journal referees.

- Garcia-Carpintero. (2019). "Semantics of Fictional Names." *Teorema: International Journal of Philosophy*. 38 (2): 73-100.
- 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): 613-635.
- Kripke, Saul A. (1980a). *Naming and Necessity*. Cambridge, MA: Harvard University Press.
 ----- . (1973b, ms). Reference and Existence: The John Locke Lectures for 1973.
- Lewis, David. (1986a). *On the Plurality of Worlds*. Blackwell: Oxford.
 ----- . (1978b). "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.
- Savage, Heidi. (2020). "How I Stopped Worrying and Started Loving 'Sherlock Holmes': A Reply to Garcia-Carpintero." *Teorema: International Journal of Philosophy* 39(1):
- Stich, Stephen (1971). "What Every Speaker Knows." *Philosophical Review* 80: 476-96.
- 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