THE MEANING OF PROPER NAMES AND THE FOURTH WALL

An Honors Thesis

Presented by:

Deniz Satik

Completion Date:

December 2014

Approved by:

Kevin Klement, Philosophy

Gary Hardegree, Philosophy
Abstract

Title: The Meaning of Proper Names and the Fourth Wall
Author: Deniz Satik
Thesis/Project Type: Independent Honors Thesis
Approved by: Kevin Klement, Philosophy
Approved by: Gary Hardegree, Philosophy

This thesis is an exploration of how the most popular theory on the meaning of proper names, direct reference theory or Millianism, handles the intuitive truth of sentences containing the proper names of fictional characters, and especially how they handle the “fourth wall” phenomenon, in which fictional characters speak directly to their audience. There are four main views proposed by Millians that are used to account for the intuitive truth of sentences containing the proper names of fictional characters. I use sentences such as “Holmes spoke to his readers” to show how each view fails to account for the intuitive truth of these sentences. The “fourth wall” sentence implies direct causation between a fictional character and a concrete, actual object, and this cannot be explained on any of the views Millians present. I conclude that, if the Millian can come up with no solution to the “fourth wall” phenomenon, then perhaps we should abandon the direct reference theory and take a different view of the meaning of proper names, such as internalism, in which there is no relationship between the meaning of words and the external world.
1. Introduction

Discourse involving fictional characters has remained a source of confusion and difficulty for various views held by philosophers of language and metaphysicians. Though numerous solutions have been proposed in the literature, philosophers have not considered, in detail, certain peculiar facts about fictional discourse which are intuitively true, and the consequences that arise for the views they hold. In this paper, I present some facts that are worth considering, and focus on one particular phenomenon known as “breaking the fourth wall.” My goal in this paper is to undermine these proposed solutions for truths in fictional discourse.

In fictional contexts and in ordinary conversation, people often utter sentences such as “Harry Potter is a wizard” which they take to be undoubtedly true, despite being aware of the fact that Harry Potter doesn’t exist. If Harry Potter doesn’t exist, then why aren’t these sentences meaningless?

There are two ways to evaluate sentences like “Harry Potter is a wizard.” The first is what Gareth Evans\(^1\) dubs as the conniving use of such sentences, in which the sentences have merely fictional truth-conditions, where they are true merely from the standpoint of the fiction. On the conniving use, the utterer is engaged in pretense when he or she utters that “Harry Potter survived the Avada Kedavra curse,” as this has a fictional truth-value and not an actual one. Another use is called the non-conniving use of such sentences, in which the sentences have real truth-conditions. For example, if someone asked you what Harry Potter’s occupation was, according to the Harry Potter stories, to say that “Harry Potter is a wizard” has a real truth-value, and not a mere

---

fictional one. A challenge for the philosopher of language is to give us a theory of meaning in which both uses of this sentence can be true, or to give us an error theory which attempts to explain away the various intuitions that speakers have about the truth values of these kinds of sentences.

Millianism, the common sense view that the semantic value of a name is its referent, runs into an unsurprising amount of trouble when attempting to analyze sentences of this kind. We assume that a sentence is meaningful just in case every constituent is meaningful. However, the proper name “Harry Potter” has no semantic value since it refers to nothing.

Yet, direct reference theorists have come up with many attempted solutions: some Millians are possibilists in that they believe we refer to objects in other possible worlds; other Millians are Meinongians, in that they believe that there are nonexistent objects; some Millians are Artifactualists, in that they identify fictional characters with dependent abstract objects; finally, some Millians hold antirealist views of fictional characters in which they deny that there are fictional objects of any kind, and they attempt to explain away the intuitive truth values of fictional discourse via some mixture of pretense, paraphrase, story operators, Millian Descriptivism and so on.

My goal is to argue against all of these views and show that Millianism cannot solve the problem of fictional discourse. In this thesis, I present, among others, a commonly used literary device within fiction known as “direct address” or “breaking the fourth wall” within popular culture, which has not been considered in detail in the literature. It is a literary device in which a fictional character speaks directly to and acknowledges the reader. For example, Italo Calvino’s novel, If on a Winter’s Night a
Traveller, addresses the reader directly in the second person and discusses the supremacy of the reader due to being able to realize the text in his or her imagination\(^2\). I use this literary device, along with many other counterexamples, to argue against all of the aforementioned views.

For the sake of simplicity, since Holmes is perhaps the most discussed fictional character in the philosophy of language, let us assume that there is a story written by Conan Doyle in which Holmes speaks directly to his reader after solving a case. Assume he says something such as “I hope you enjoyed this book!” after looking up at the ceiling. The story could even be based on an investigation on the identity of the reader. Such a story could easily exist; if necessary, we could write ourselves an actual novel in which this occurs, and readers can believe it to be true according to the story.

Now, let’s consider the following sentence:

1) Holmes spoke to his readers.

This sentence seems as intuitively true as the sentence “Holmes spoke to Watson.” A prima facie reason why this sentence is so troubling is that it seems to imply that Holmes, who doesn’t exist, can interact with and speak to his readers. However, from this implication, you can infer that Holmes does exist, which seems to contradict our intuition that he doesn’t exist.

In this thesis, after presenting each of the Millian solutions, my aim is to use the first sentence, along with variations of it and other specific counterexamples, to argue against all of these views. My goal is to show that none of these views have the specific ontological requirements to solve this problem, and any attempt to paraphrase “fourth

---

wall” sentences would be ad hoc or inelegant.

After doing so, I want to consider the implications of the “fourth wall” phenomenon. What kind of a solution can the direct reference theorist propose? Can they give up the view that fictional characters have proper names, or can they deny that “direct address” events actually occur? Finally, I suggest that, if the Millian can provide no solution for this problem, then perhaps we should give up Millianism. Only then we may be able to account for the intuitive truth of certain sentences involving fictional characters.

2. Possibilism

Possibilism is the view that fictional entities are entities that exist not in the actual world, but in other possible worlds. On this view, there are possible worlds in which the Sherlock Holmes stories really do happen, and Holmes really does solve cases and defeat Moriarty. Perhaps the most well-known and developed of this kind is in Lewis (1978), which is successful at solving many of the problems posed to it by philosophers of language such as Kripke in Naming and Necessity.

Lewis thinks that we should not take the non-conniving use of fictional sentences such as “Holmes is a detective” at face value; rather, he believes that we should paraphrase them with an intensional operator such as “according to the Holmes stories.” The sentence with the operator is true just in case Holmes is a detective in the worlds picked out by the intensional operator. The sentence “Holmes is a detective”

---

3 This view immediately runs into the problem of impossible objects and stories. For example, there are many stories in which all sorts of laws of logic are broken: there are stories with square circles, stories in which the protagonist travels his time, kills his grandfather and continues to exist, and so on. We might want to add in impossible worlds to account for the truth of sentences with such impossible events, but this comes at the cost of triviality. More work on the metaphysics of impossible worlds would shed some light on this matter.

without the operator is either false or meaningless\(^5\), depending on one’s view.

Lewis does not think it is right for the operator to just pick out all the worlds in which all of the events as described by the story are true. Let us assume that Conan Doyle wrote the stories as fiction. As Kripke points out, by pure chance, it may be that our own world could be a world in which the Holmes stories take place without Doyle knowing of it.\(^6\) If Sherlock Holmes happened to have existed in real life without Doyle’s knowledge, this Holmes still would not have been the Holmes that Doyle wrote about. After all, it is false that in our world, the name “Sherlock Holmes” refers to someone though it is true in the stories. He does not exist.

Lewis’s solution for this is as follows. We should consider the worlds in which events are qualitatively identical to the events as described in the fiction, in which the fiction is told as a story told by a particular storyteller at some time, about known facts rather than about made up stories. Such a story must be told by Conan Doyle himself rather than anyone else. For example, if I were to tell the same story word for word, that would not make it the same story as Conan Doyle’s. Different acts of storytelling lead to different stories. In these worlds, the proper name “Sherlock Holmes” behaves as a rigid designator, though it does not in our world. So it cannot designate anyone in our world: hence, it can account for Kripke’s problem that we cannot identify a fictional object with a possible object.

This analysis still has a few problems. It brings in far too many worlds in which all sorts of things are true. It is intuitively true, but never stated, that Holmes is not an alien from Mars, and that Holmes does not eat mud, and that Holmes has to use the

\(^5\) This is because there is no actual object corresponding to the fictional character “Holmes.” He is only a possible object.

restroom often. But this analysis brings in all these worlds in which the opposite is the case. Lewis suggests that we analyze true fictional statements as counterfactuals, and pick out the worlds which differ the least from our own world.

We pick out the worlds which differ the least from our own world, but which world is the actual world? It is a contingent matter which possible world is the actual world. There are still many things we don’t know about the actual world: it could be possible that the Holmes stories have truths which no one knows about. Lewis suggests that, instead of using the actual world as a standpoint for our fictional truths, we should use the worlds in which the overt beliefs of the community of origin are true—the community of origin being the community in which Doyle wrote the stories.

Lewis’s analysis solves many well-known problems that plague possibilist accounts. For example, Kripke argues that there are too many possible worlds to pick out—there are worlds in which Holmes is from Tatooine and worlds in which he eats maggots. As Lewis says, after restricting the possible worlds we can consider, there’s no point in picking out the world in which the stories take place. There are worlds in which Holmes has the blood type A and others in which he has the blood type AB—it is absurd to suppose that the story has answers for these. The worlds of the Holmes stories are plural, and different worlds may have different answers to these questions.

A serious problem worth pointing out with Lewis’s view is that he offers no account of the truth of sentences such as “Holmes is a fictional character.”

Regardless, it should be clear that the “fourth wall” objection is damaging to this account.

---

7 Also consider examples such as the following: “Holmes is smarter than Gregory House”; “Holmes appears in more than one literary work.” Holmes is identified with a possible object, and in the first sentence, we’re comparing possible objects from different possible worlds with each other. We cannot do this with the intensional operator. In the second sentence, the sentence is true in the actual world and it is not true according to the story. And we cannot identify Holmes with any actual object on this view.
1) Holmes spoke to his actual readers.

We assumed earlier that this sentence is true according to the story, as Doyle wrote it. So we can read this with the intensional operator as Lewis suggests:

1a) According to the Holmes stories, Holmes spoke to his actual readers.

On Lewis’s view, it is hard to make sense of the truth of this sentence. It says that Holmes, a possible object, spoke to his actual readers, which exist in this world. This is impossible: possible worlds do not have any sort of spatiotemporal relation with each other and Lewis’s counterfactual analysis of causation prevents the possibility of trans-world causation.\(^8\) It may be impossible for this case of trans-world interaction to be true.

One difficult way out of this might be to suggest that “speaking to” is not a causal relationship between two objects.\(^9\) It may be a mere intentional one, as seen by sentences such as “Gary spoke to his intro to logic students, but none of them listened.”\(^10\) The sentence denies that there is any kind of direct causal relation between Gary and his intro to logic students but it still is intuitively true. So in sentence 1, Holmes does not have a direct causal relationship with us, and it is not necessary for this sentence to be true. But suppose we had a version of sentence 1 in which we could force the causal reading between Holmes’s action in another possible world and us in the actual world:

1b) When Holmes spoke to his actual readers\(^{11}\), he made some of them laugh.

This sentence seems intuitively true as well. What this sentence says with the

---


\(^{9}\) Many thanks to Kevin Klement (personal communication) for this objection.

\(^{10}\) This potential counterargument for the “fourth wall” objection only exists when we consider possible objects as concrete objects as Lewis does, in which case they can have intentional states. However, if we take the much more common view of possible worlds as abstract, this cannot be used as a counterargument, because on the abstract view of possible worlds, possible worlds are merely states of affairs and abstract objects cannot have intentional states.

\(^{11}\) The first embedded sentence is read with the intensional operator: the entire sentence as a whole is not, and cannot be read with the intensional operator. So when according to the story, Holmes spoke to his actual readers, some of his actual readers laughed.
intensional operator is this: according to the story, when Holmes spoke to his actual readers from the world of fiction, his action was such that it caused some of his actual readers to laugh. We are back to where we started, since this sentence implies causation across worlds.

In that case, suppose that one objected as follows: the sentence is intuitively false, because when we read the event as described in the story and laugh, it’s not really Holmes that makes us laugh. It’s thinking about the event as described in the story that makes us laugh, and not Holmes interacting with us across possible worlds that does. The problem with this argument is that there seems to be a sense in which fictional characters really can make their audiences react in certain ways.

For example, it is completely felicitous to say “Holmes pissed me off in Doyle’s latest book” or “Holmes made me happy when he kissed Watson on the lips.” Such sentences are not directly causal, for in them, Holmes is not even attempting to interact with the reader. This line of reasoning holds true in real life sentences such as “Obama made me furious with his latest immigration bill.” Obama does not have a direct causal relationship with the utterer, because what makes the utterer mad is thinking about Obama’s latest immigration bill, and not Obama interacting with the utterer.

But when we say “Obama made me furious when he wrote to me,” we now understand that there is a direct causal relationship between Obama and the utterer. So why can’t we apply this intuition to fictional characters, or in other words, why should we draw the line there? It seems to be natural and intuitive to think of sentence 1b as a causal relationship between a fictional character and the reader, as there is now direct interaction between Holmes, and the reader, in the same way that there is direct
interaction between Obama and the utterer. So there doesn't seem to be any reason to deny the intuition that sentence 1b is true.

Lewis may attempt to provide another analysis of such sentences, in which they are true just in case Holmes really does speak to a future reader or an observer of his real life endeavors in the worlds of fiction. But what we care about in sentence 1 is that Holmes spoke to his readers in the actual world, and not in any other possible world.

Indeed, Lewis's account of fictional characters runs into trouble when dealing with metafictional truths of any kind, in which fictional characters are aware of the fact that they are fictional characters, and numerous other facts that follow. Suppose the sentences below are true according to the story:

2) Holmes knows that he does not exist.

3) Holmes knows that he was created by Conan Doyle.

We can attribute beliefs of this kind to any fictional character, and speakers can believe them to be true. On Lewis's view, possible worlds are just as real as the actual world, and they are no different in kind from the actual world. It would be contradictory for Holmes to think that he does not exist if this is the case. Furthermore, in the worlds where such stories are told as known fact, Holmes can at most only be aware of the fact that his story was told by Conan Doyle to other people. It would be strange for him to think that he was literally \textit{created} by Conan Doyle.

To be fair, there is nothing logically impossible about having these beliefs. Indeed, Lewis can even allege that Holmes may be deluded and have these beliefs. But it cannot be true for Holmes as it is for us: from the standpoint of the story, we know that Holmes is certainly not deluded (since the story says so), and he can be justified in
having these beliefs as a fictional character. We cannot say that Holmes is justified in having such beliefs if we think that stories of fiction take place in other possible worlds. So possible worlds are not a good metaphysical foundation for fictional discourse.

Finally, we can also consider sentences from the standpoint of the Holmes stories. We can have an utterance in the story as follows:

4) I do not actually exist.

According to Lewis’s analysis of actuality, it is an indexical whose extension is determined by the context of the speaker’s utterance.\(^\text{12}\) In other words, it would pick out the world in which Holmes uttered this sentence. Let’s analyze this sentence from Holmes’s perspective. As he is a mere fictional character, we expect it to be true. But this sentence is self-contradictory on Lewis’s view, although we can make perfect sense of it from Holmes’s standpoint and agree with Holmes if he were to utter this sentence.

While Lewis’s view is successful at solving many of the generic possibilist problems, he also runs into severe issues with metafictional sentences, in which fictional characters can be aware of the fact that they are fictional characters, and can speak to their actual readers, among other issues that have been brought up by Lewis himself. The truth of perhaps the most obvious sentence involving the name of a fictional character, “Holmes is a fictional character,” is not given any kind of account on this view. So possibilist accounts not only have trouble accounting for the truth of metafictional sentences, but they also don’t provide any kind of account for the actual, real-world truth of sentences with fictional names. We will now consider other views.

3. Meinongianism

Meinongians distinguish between the meanings of “there is an x” and “x exists.” To say that there is a square circle is to say that there is an x such that x is a square circle, while to say that a square circle exists is to say that there is an x such that x exists and x is a square circle.¹³ For the Meinongian, then, there are nonexistent objects. Meinongians disagree about what kinds of nonexistent objects there are: for example, Zalta (1983) believes that there are two kinds of predications, and that Meinongian objects are abstract objects, whereas Parsons (1980) believes that there are two kinds of properties, and that Meinongian objects are concrete objects. We will consider both views.

According to Parsons, we must distinguish between two kinds of properties.¹⁴ There are nuclear properties, which are the normal properties we attribute to concrete objects such as being red, tall, blue or thin. We also have extranuclear properties, of which we have four kinds: ontological (such as to exist), modal (to be possible, necessary), technical (to be complete) and intentional (to be thought of by others).

For example, the character Sherlock Holmes has the nuclear properties of being a detective, of being a flesh-and-blood human being, while he has the extranuclear properties of being a fictional character and not existing. Crucially, sentences such as “Holmes is a detective” and “Holmes is a fictional character” are both as true and read straightforwardly without the use of an operator.

As Parsons points out, while impossible objects may be problematic for the

---

possibilist view, they are not problematic at all for his. For example, we can say that a
square circle is round and not round at the same time. On Parsons's view, this just
means that square circles necessarily cannot exist. The principle of non-contradiction
does not need to hold for nonexistent objects. Nor do nonexistent objects have to be
complete: an existent object must be P or not P for every predicate P, but for Holmes,
the nonexistent object, we surely cannot say how many pairs of underwear he owns.

Zalta (1983), on the other hand, believes that Meinongian objects are abstract
objects for which there are two kinds of predications: exemplifying and encoding. For
example, the man Edward Zalta can exemplify the properties of being a human being
and being a professor. But the abstract object, Holmes, cannot exemplify these kinds of
properties: he merely encodes, or describes the property of being a detective. He
encodes all the properties which the story says he has. For Zalta, abstract objects can
also exemplify certain properties that Parsons defines as “extranuclear.” The abstract
object Holmes can exemplify the property of being a fictional character.

Zalta’s view, just like Parsons’s theory, can solve the problems posed to
Meinongians. We do not have to give up the principle of non-contradiction since a
square circle merely encodes the properties of being round and non-round, and the
principle only applies to exemplification of these properties. Furthermore, while every
concrete object must exemplify, for every predicate P, either P or not P, abstract objects
may be incomplete in terms of the properties they encode.

As we can see, both views handle sentences of the kind “Holmes is a detective”
and “Holmes is a fictional character” very well. However, it is important to note that
Parsons and Zalta have a very different analysis of the first sentence: for Parsons, the
nonexistent object exemplifies the property of being a detective, whereas for Zalta, the abstract object Holmes merely encodes it. In my opinion, Parsons’s view may be more in line with our intuitions: for example, that “Harry Potter and Deniz are both human beings” seems to be true, and to say that “Harry Potter is braver than any actual man” seems to be felicitous. Zalta’s predications would not allow us to make these comparisons, since abstract objects do not exemplify the property of being “human” or being “brave.” They merely encode these properties.

How do both views handle the “fourth wall” objection?

1) Holmes spoke to his readers.

The problem is simple: the sentence implies that Holmes communicated and interacted with existent objects. It would be peculiar and impossible for Holmes as a nonexistent object to do this, because nonexistent objects cannot have any direct causal relations with existent ones. They are just two fundamentally different kinds of being.

There is a sense in which Meinongian objects can have causal relations with us. For example, to say that “square circles piss me off” can be true in certain contexts in which the utterer really does get angry thinking about square circles. But this is a different sense of causation than the causation in the “fourth wall” sentences. The “fourth wall” sentence says that a nonexistent object, Holmes, is interacting directly with the reader. But the aforementioned sentence merely says that thinking of the idea of a square circle makes the utterer angry. There is no direct causal relationship between the square circle and the person, since the square circle is merely being thought of. The square circle is not actively interacting with the person thinking of the square circle.

The Meinongian can also come up with the same counterarguments as the
possibilist, but such attempts would fail for the reasons we have considered earlier. So the Meinongian has trouble telling us why this sentence is true. An advantage for the Meinongian, however, is that they do not have any trouble in accounting for the truth of the aforementioned sentences we presented against the possibilist view.

It is not impossible for the Meinongian to solve this problem, though it comes at a great cost. For sentences of this kind, we can paraphrase them with a story operator, which would mean that the sentence is true just in case the story says it is true. Even if the embedded sentence implies something that is absurd, it is no problem as long as it is true according to the story.

1p) According to the Holmes stories, Holmes spoke to his readers. This is a context-shifting operator. It is a two-place function from the context of the embedded sentence and the chosen story that returns as a value the context of the story, the indirect context. So 1p will be considered in an indirect context, and the truth of the sentence will be accounted for, as the embedded sentence is true according to the story.

Thomasson (1998), whose artifactualist view we will shortly consider, is very critical of such maneuvers solely to avoid reference to objects. To see her point, we can make our own example:

5) Holmes spoke to Watson. This sentence is surely true as Holmes has spoken to Watson many times during the Holmes stories. And as we said earlier, such sentences which are true according to the story are read without the operator in both of the Meinongian theories we have

---

16 Notice that on Zalta's view, abstract objects cannot have intentional states, so that objection would not work regardless.

17 It is worthwhile to note that such a move faces a well-known criticism by David Kaplan in *Demonstratives*, who dubs such operators as “monsters.” But for the purposes of the argument, we will, just for the time being, grant the Meinongian this move.

considered. Why is sentence 5 read straightforwardly by speakers, whereas sentence 1 is awkwardly read with the operator? Intuitively, it’s clear that both sentences are read without further thought. As Thomasson argues, we must read these two sentences differently solely because they cause trouble for the Meinongian even though they are taken in the same context with the same surface grammar.

This is not the only context which the Meinongian must paraphrase. Thomasson’s own objection to Meinongianism is very similar in that it accuses the Meinongian of relying on paraphrase. Let us say that there is a story in which George W. Bush became a cowboy instead of becoming the President of the United States.

6) Bush is a cowboy.

As Thomasson argues, it is an empirical fact that such a sentence is false, though on the Meinongian view it is read straightforwardly without the operator. Of course, as Parsons claims, one could postulate that there is a fictional surrogate of Bush which is a nonexistent object. But on this view, a fictional story can never be about a real person, place or an event. For this reason, Parsons himself does not favor this view, for intuitively, the real object, George Bush, is the one who appears in the story.

The Meinongian may opt to paraphrase 6 in the same way they paraphrased 1, but this is false for the reasons we considered earlier. The sentence “Indiana Jones is a cowboy” is read by straightforwardly by speakers, as would “Bush is a cowboy” in certain contexts. The Meinongian cannot explain why speakers have the same intuitions about both sentences even though they involve different kinds of objects.

We now have two kinds of sentences that must be paraphrased solely because they have references to objects that cause trouble for the Meinongian, leading to a very
inelegant and awkward theory. So the Meinongian seems to have a great deal of trouble accounting for truths in fictional discourse. We will now consider Thomasson’s own view, which is perhaps the best analysis of fictional discourse.

4. Artifactualism

Zalta’s and Parsons’s views of Meinongian objects have the additional problem of being unable to account for the idea that fictional entities come into being once their authors conceive of and write a story about them.\(^{18}\) For the Meinongian, there is a nonexistent object for every set of properties. So the author is only responsible for picking a certain set of properties when coming up with a character. But this seems to contradict the intuition that authors do not merely select a certain fictional entity.

On Thomasson’s account of fictional entities, they are abstract objects, objects which don’t exist in space-time, which come into being after their creator conceives of the entity and places the entity into a certain narrative. On her view, it is necessarily true that for a fictional entity, \(X\), created at time \(t\), the author exists at some time before \(t\), \(t’\), and it is also necessarily true that if \(X\) exists, then a fictional work featuring \(X\) exists. This allows us to account for the intuition that fictional characters are created by the author of the story they are in. This is, however, at odds with the idea that abstract objects exist eternally, which may be an issue for a metaphysician.\(^{19}\)

She suggests that we consider two different contexts in which fictional discourse takes place: there is the fictional context, which is identical to what Evans dubs the conniving use, in which speakers pretend that what happens in the story is true, so


\(^{19}\) For example, it is not clear how an abstract object, which does not exist in space-time, can be created at a time \(t\), which implies that it does exist in space-time.
sentences such as “Holmes is a detective” are taken straightforwardly and read without the operator. The use of pretense renders the use of such operators unnecessary.

When we move on to the real context—what Evans dubs the non-conniving usage—we can talk about in fictional characters in two ways. The first is in which we utter truths about fictional characters that are not true merely according to the story: for example, that Holmes is in many literary works is true even though the story does not elaborate on how many works he is in. The sentence is true from our perspective and not from the perspective of the story.

Thomasson’s view can account for these sentences. Her abstract objects exemplify many of the extranuclear properties that Parsons defines: for example, abstract objects can be fictional characters, appear in more than one literary work, and be detectives according to the Holmes stories and so on. So sentences such as “Holmes is a fictional character” come out as true and are read straightforwardly, with no operator. But, for example, to say that “Holmes does not exist” is false, on her view, despite it being intuitively true. These abstract objects, unlike on Zalta’s view, are not nonexistent objects.

And we also have truths according to the story, for example, that Holmes is a detective. For the latter, Thomasson suggests that we read all kinds of such sentences with the aforementioned story operator that the Meinongian took advantage of. The latter sentence would be true just in case the story says he is a detective. And it does. Thomasson, just like the Fregean, puts the sentence in an indirect context, so that the sentence can still have a truth value. Unlike the Meinongian, she does not take the embedded sentence to be actually true.
This allows us to avoid the problems that the Meinongian suffers from. For example, while the sentence “Bush is a cowboy” is not true in the real world, the sentence “according to the Bush stories, Bush is a cowboy” is true in the stories about George Bush being a cowboy. This allows us to include the intuition that actual people can appear in fictional works, and avoid postulating fictional surrogates.

Earlier, we brought up the unintuitive conclusion we obtain from Thomasson’s theory that Holmes exists. While this sentence would be false according to the average speaker, for the artifactualist, unlike Zalta (1983), abstract objects do exist, so this sentence would be true. She thinks we should paraphrase such sentences as follows: Holmes doesn’t exist (as a concrete object). Abstract and concrete objects certainly don’t exist in the same sense, even if both do ultimately exist in one way or another.

That sentences relating to truths from the stories are read with operators is a great advantage for the artifactualist. Let us see how she might account for the truth of the “fourth wall” example:

1p) According to the Holmes stories, Holmes spoke to his readers.

We explained how the story operator works earlier. The embedded sentence, “Holmes spoke to his readers” is put in an indirect context so it does not matter if “Holmes” refers to an abstract object which cannot speak to its readers. So the sentence comes out as true, as intended. At first glance, this sentence poses no problem for the artifactualist. And one cannot accuse Thomasson of being ad hoc.

For Thomasson claims that the operator is a part of the surface grammar of sentence 1, and every sentence in this context is read with the operator. The Meinongian, on the other hand, wishes to only use the operator in certain troublesome
contexts. Thomasson’s theory is, in this way, much more elegant than the Meinongian’s. But as we shall see, there is a sense in which the sentence is not merely true according to the story, but it is also actually true.

Perhaps the major problem with this view is that it fails to account for the intuition that fictional objects actually have the properties that the book ascribes to them. Lewis (1978), for example, believes that there is a sense in which Holmes is a real life flesh-and-blood person just like Richard Nixon. Lewis’s modal realism can account for his intuition, since Holmes exists in another possible world just as concrete and as real as ours. However, it runs into trouble when we attempt to compare entities from different possible worlds with each other.

Further, Kroon and Voltolini (2011) argue that artifactualism fails to account for the idea that, as abstract objects, fictional objects actually have the properties that the story ascribes to them. They consider the following example:

7) Holmes is cleverer than any actual detective.

In this sentence, we seem to be saying that Holmes actually has the property of being smart, and is not merely smart according to the story, as Thomasson’s view would claim. There are many more sentences which allow us to see this point:

8) Holmes and Quine are both human beings.

This sentence is intuitively true, but on Thomasson’s view, Holmes and Quine do not share this property in common, since Holmes is a human being merely according to the story while Quine really is a human being.

9) Holmes and Quine both spoke to their readers.

---

As relations such as “x spoke to y” are nothing more than two-place properties, we expect this property to be actually true of Holmes just like the others. Let’s assume there’s an event in which Quine spoke to a reader of his papers in some way, perhaps saying “I hope you like my rejection of the analytic/synthetic distinction!” In that case, this sentence would again be intuitively true, since both Holmes, the fictional character, and Quine, the philosopher, spoke to their own readers. The truth of this sentence and the others makes it seem possible for us to actually attribute any kind of property that is true according to the story, to a fictional character.

As we can see, the Meinongian has the advantage here: Meinongians can account for the truth of two of these sentences—but not the third—since nonexistent objects really do have the properties that the book ascribes to them. Now, let’s go back to the “fourth wall” objection:

1) Holmes spoke to his readers.

According to Thomasson’s view, Holmes only possesses the property of speaking to his readers according to the Holmes stories. So her theory cannot account for why this sentence is actually true: after all, a mere abstract object like Holmes which does not exist in space and time cannot interact with and speak to concrete objects. Abstract objects have no causal relations with concrete objects, as they do not exist in space and time, and they do not have intentional states.

Any sentence such as “Holmes is a detective” which can be actually true will suffice for this objection. We do not need the “fourth wall” sentence to object to Thomasson’s view. However, the “fourth wall” sentence is necessary as a counterexample to a certain kind of Artifactualism, which attempts to combine the most
promising elements of Meinongianism and Artifactualism to solve the problem of sentences such as “Holmes is a detective” being actually true.\textsuperscript{22} For such a view, the “fourth wall” sentence is equally damaging.

While the Artifactualist theory can explain many truths that other theories run into problems with, identifying fictional characters as abstract objects that cannot have nuclear properties comes at a great cost. There are many intuitively true sentences for which an account can only be given by admitting that fictional objects really do have the properties that the story attributes to them.

Therefore, Artifactualism cannot account for why the “fourth wall” sentence is actually true, even if it can account for it being true according to the story, unlike the other two views we have seen.

5. Antirealism

As we have seen, it may be impossible to find the right kind of fictional entity which can account for all of the intuitions that we have about fictional characters. For example, Russell (1919) beautifully proclaims\textsuperscript{23}:

\begin{quote}
Logic, I should maintain, must no more admit a unicorn than zoology can; for logic is concerned with the real world just as truly as zoology, though with its more abstract and general features. To say that unicorns have an existence in heraldry, or in literature, or in imagination, is a most pitiful and paltry evasion.
\end{quote}

In line with this reasoning, and in order to maintain a scientifically elegant view with minimal ontological commitments, many philosophers of language have wisely attempted to build a theory of meaning about the nature of fictional discourse with no


reference to objects. Among Millians, Walton’s pretense theory is perhaps the most popular antirealist view.

As Thomasson admits, pretense is surely a part of fictional discourse in certain “fictional contexts,” or in their conniving usage, in which speakers pretend as if what the story says is true. In such a context, saying that “Holmes is a detective” has a merely fictional truth-value, and it is understood by the speakers that some form of pretense is at play. Walton believes that we can extend the use of pretense to the non-conniving usage of such sentences, in which there is merely a different kind of pretense.24

On Walton’s view, when we use sentences in their fictional context, we engage in a pretense of a certain kind in an authorized game. For example, the sentence “Tom Sawyer attended his own funeral” would be understood as follows:25

*The Adventures of Tom Sawyer* is such that one who engages in pretense of kind K [claiming “Tom Sawyer attended his own funeral”] in a game authorized for it makes it fictional of himself in that game that he speaks truly.

To engage in “pretense of kind K” simply means to make a certain claim, such as “Holmes is a detective.” The stories serve as the background for what goes on in this authorized game, which makes it inappropriate for us to utter sentences such as “Holmes is not intelligent.”

In the real context, we engage in a pretense of a certain kind in an unofficial game. Unofficial games are not necessarily related to authorized games, and they can exist without a fictional background. However, the background information can still determine what is apparently true and what is not in certain sentences in the real context. In an unofficial game, the apparent truth of a sentence such as “Holmes is a

---


detective‖ in its real context depends on its being appropriate in an authorized game.

For example, on Walton’s view, to say that “Gregor Samsa is a fictional character” is to say the following:

There may be an unofficial game in which one who says [“Gregor Samsa is a fictional character”] fictionally speaks the truth, a game in which it is fictional that there are two kinds of people: “Real” people and “fictional characters.”

This is noticeably awkward. To avoid reference to fictional entities, Walton’s pretense theory forces us to think that every statement containing the proper name of a fictional character must have some sort of pretense; that is, we seem to be pretending (without knowing that we are) that Holmes is a fictional character, and that Holmes appears in more than one literary work, and that Holmes is a detective according to the stories, and so on. This doesn’t seem to be very representative of fictional discourse. It seems to be an empirical fact that Holmes does appear in more than one literary work rather than a mere matter of pretense that we’re not even aware of.

Walton’s theory can account for the “fourth wall” objection as presented:

1) Holmes spoke to his readers.

On Walton’s view, this sentence is false. After all, according to Walton’s pretense theory, what makes this sentence appear true is that we simply pretend as if this sentence is true with the book as a prop, dictating what is true and what is not. As opposed to other views we have considered, Walton’s pretense theory provides us an error theory for fictional discourse: it is not really the case that Holmes spoke to his readers and that he is a detective. It is only a matter of pretense.

Let us update the first sentence:

1a) According to the Holmes stories, Holmes spoke to his readers. On Walton’s view, this sentence is again false and we must pretend as if it’s true due to an unofficial game. Any other sentence such as “According to the Holmes stories, Holmes is a detective” would again be false when it seems just as true as to say any other empirical fact. Walton’s maneuvers, then, make us reject obvious empirical truths, and make it seem possible for us to pretend something without knowing it, merely to avoid reference to ontological objects. Let’s build on the latter objection.

There are more counterexamples worth considering to the idea that we can pretend something without knowing that we are. We can imagine a child, or a person, who thinks that Pokémon like Pikachu or people like Harry Potter or Santa Claus really exist—which is true of many children today—and we can say they uttered the following intuitively true sentence:

10) Harry Potter is actually famous.

But we have already assumed that this child is not pretending in any sort of way that this sentence is true, as he or she thinks that Harry Potter really does exist. So why is an utterance of this sentence by this child true without any sort of pretense at play? When questioned, the child may indignantly continue to insist that Harry Potter is famous in the real world, and it is a genuine empirical fact that he is. So it seems unrealistic to think that a child can pretend without knowing that he or she is pretending.

Walton’s theory also requires some ad hoc maneuvers in order to account for the truths of certain sentences:

11) Holmes and Quine both spoke to their readers.

12) Russell and Quine both spoke to their readers.
On Walton’s view, 12 is a true empirical fact, and read completely straightforwardly. On the other hand, we need various complicated ad hoc paraphrases in order to make sense of 11. The apparent truth of the first conjunct of 11, “Holmes spoke to his readers,” depends on its being true in its fictional context, which it is. And we then need to come up with another game to explain why sentence 11 is true. But 11 and 12 both seem to be equally straightforward and easy to understand.

To sum up, Walton’s pretense theory really is successful at showing us how pretense takes place in certain contexts. But the main problem is that his view forces us to take every sentence containing the name of a fictional character as a form of pretense, which seems profoundly unintuitive. His view also makes the intuitively wrong psychological claim that we can pretend something without knowing that we are, even about obvious empirical facts like “Holmes is a fictional character.”

6. Conclusion

We have seen that solutions proposed by Millians have a lot of trouble accounting for the truth of certain sentences. For Lewis’s view, it implied trans-world interaction, which is impossible. For the Meinongian, the view implied that nonexistent objects can talk to existent objects, and attempts to fix this turned out to be very ad hoc. The Artifactualist ran into the problem of accepting that sentences which are true according to the story can also actually be true. Finally, the pretense theory required major ad hoc adjustments in order to avoid reference to any sort of fictional entity, and it seemed out of line with how fictional discourse takes place.

After being convinced by such arguments, the direct reference theorist may wish to deny that the names of fictional characters are true proper names, and that the theory
of direct reference is therefore silent on whether the names of fictional characters refer to entities at all. What could distinguish the name “Holmes” is that it is not being used in an attempt to refer to an object by its users, which is what prevents it from being a genuine proper name.

There are many reasons to believe that the names of fictional characters are no different. It seems that no matter what language you look at, from English to Turkish, you can replace the name of a fictional character with the name of a real person in any sentence and the sentence would still be syntactically well-formed, though it may have an awkward meaning depending on the content of the sentence. The proper names of fictional characters seem to behave the same, syntactically, as ordinary proper names. To say that “Holmes is a detective” makes us wonder whether it is true or false to attribute the property of being a detective to the bearer of the name “Holmes,” just like other sentences do.

Second, it would be impossible to account for the utterances of a child who believes that Pikachu, Harry Potter or Santa Claus are real. The child who doesn’t know that Pikachu doesn’t exist uses the name intending to refer to an object that exists. On this view, the name of a fictional character like Pikachu is a genuine proper name. So Pikachu is not a fictional character. Yet this is not the desired conclusion. Pikachu is clearly a fictional character.

Another option would be to deny that there is such a phenomenon such as “direct address” or “breaking the fourth wall” at all. One could think that these are merely literary devices used by the author to address the reader, and they are not really part of the story, or true in the established facts of the fictional universe.
This is true to an extent. For example, there are video games in which the main character tells the player which buttons to press, and the player is never mentioned again during the story. In that case, a “fourth wall” sentence would intuitively come out as false, as it’s a mere tool for the author to communicate with the player. But to say that “direct address” is never true in a fictional universe would be going too far.

In fiction, whatever the author intends to be true according to the story is the case according to the story. Suppose the author of the Holmes stories intended for “fourth wall” events to be true according to the story, and the readers accepted such events as canon, or in other words, as a part of the fictional universe. We can imagine the Holmes stories being based on an investigation to figure out who the reader is, or on entertaining the reader. In that case, “breaking the fourth wall” is surely not a mere literary device used by the author, and it is a genuine part of the story and the established facts of the fictional universe.

In explaining the truths of fictional discourse, the way of the Millian is difficult. Therefore, a view of proper names which avoids the necessity of such ontologically desperate and ad hoc maneuvers—in particular, an internalist view on the meaning of proper names as defended by Paul Pietroski—may be a preferable alternative to Millianism.

---


28 And it’s unlikely that a descriptivist view of proper names can handle the “fourth wall” sentences as presented here, as descriptivist views on fictional sentences use paraphrase and pretense to get such sentences to come out as true. A descriptivist account of sentences containing the names of fictional characters is very similar to an Artifactualist or an Antirealist account of such sentences. And we have already seen why such maneuvers do not work, but this is another topic for another paper.
Bibliography


