A REPRESENTATIONAL ACCOUNT OF FICTIONAL CHARACTERS

ZOLTÁN VECSEY
vecseyz@freemail.hu
MTA-DE Research Group for Theoretical Linguistics

ABSTRACT. According to the artefactual theory of literature, fictional characters are contingently existing abstract objects. Sherlock Holmes, Fyodor Karamazov and other creatures of fiction are contingent existents because they are brought into being by the creative acts of their authors, and they are abstract objects because they are not denizens of the spatiotemporal world. Although the artefactual theory seems to correspond to our literary practices, it has some counterintuitive features. The paper will propose a modified and more plausible version of artefactualism. The basic idea is that fictional characters must be thought of as linguistic representations. After analysing the representation-dependent features of authorial creation, I will argue that the proposed view helps us to make progress in the debate on the ontological status of characters.

Keywords: fictional characters; artefactual theory; abstract objects; non-relational representation


Received 5 January 2015 • Received in revised form 23 April 2015
Accepted 24 April 2015 • Available online 1 November 2015

1. Artefactual Theories of Fictional Characters

One comparative advantage of artefactual theories over their rivals is their conformity with our pre-theoretic intuition according to which fictional characters are created entities. Everyday consumers of literary works often assume that artistic writing is an essentially innovative activity. On the most common view, persons like Sherlock Holmes and Fyodor Karamazov, and places like Middlemarch and Lilliput are, in the strictest sense of the word, intellectual products of the human mind. Fictional characters are therefore supposed to have their own life span: there is a certain point of time when
they are brought into being by the creative acts of their authors and there may come a time when they simply cease to exist, perhaps, when the works in which they appear become physically destroyed and no one remembers them any more. In this regard, fictional characters do not differ significantly from other types of artistic products. It is not very surprising, then, that the folk ontogenetic conception of literary works takes the createdness of fictional persons and places as a rather obvious and trivial datum.

Most proponents of the artefactual theory tend to regard the above ontogenetic truisms of the folk view as basically correct. At least two elements of this view are thought to be suitable for serious theoretical investigations. The first element coincides with the fundamental thesis of art ontological realism which states that there are fictional characters in the overall inventory of what exists. The second element attempts to give an answer to the question of how can there be such characters. According to the folk explanation, fictional characters come into being as a result of creative writing processes.

Artefactualists such as Goodman (2004), Voltolini (2006) and Thomasson (1999, 2003, 2010) argue, in harmony with the folk view, that Sherlock Holmes, Middlemarch and their likes are contingently existing created entities. Even if this thesis appears to be reasonable at first glance, the artefactual theory has some counterintuitive features. On my view, artefactualism becomes counterintuitive precisely at that point where it goes beyond the truisms of the folk theory. Let me explain what I mean by this.1

In order to make their theory plausible, artefactualists have to provide a detailed description about the process of authorial creation. If one regards a particular fictional text $F$ as consisting of a set of syntactically individuated English sentences, then the question for artefactualists to answer is this: How can the author of $F$ literally create various fictional entities merely by writing down or typing out the sentences composing $F$?

As a first step in answering this question, artefactualists may invoke their favored account of intentionality. Thomasson (1996, 1999) follows in this respect the act-object theory of intentionality defended first by the polish philosophers Twardowski and Ingarden. The act-object theory says that every intentional act involves an object and a content. The act is directed or related to an object and the content is what is thought or judged about that object. According to this approach, non-fictional sentences like (1) receive a quite straightforward interpretation:

(1) Günter Grass smokes a pipe.

The object to which the underlying intentional act is related is Grass, and he is thought of as acting in a certain way, namely as smoking a pipe. Of course, Grass as a person exists independently from any intentional act. And
he would surely not cease to exist, if no intentional acts were directed towards him in the future. But now consider the fictional counterpart of (1):

(2) Sherlock Holmes smokes a pipe.

Suppose, for the sake of the argument, that (2) is the sentence token which contained the very first occurrence of the proper name Sherlock Holmes in Conan Doyle’s oeuvre. Presumably, when Doyle has written down this sentence, his intentional act were directed towards a putative object. But given that Holmes did not existed before (2) was written down, we may question how this was possible. What was the intended object of Doyle’s intentional act? If intentionality is essentially relational, as the act-object theory assumes, then there is a certain sort of circularity here. Doyle’s intentional act presupposed the existence of a particular act-independent object, but it seems that the presupposition in question was made by the act itself. According to Thomasson (1999), this kind of intentional circularity is merely apparent. Those who demand an act-independent object in order for (2) to have a non-circular interpretation do not take into account the creative capacity of intentional acts. We should not deny that Holmes did not exist before the tokening of (2). Prior to this time, there was no such person as Holmes and no other fictional text contained him (or it) as a protagonist. It does not follow from this, however, that (2) has to be regarded as a product of an objectless intentional act. Once we recognize that Doyle’s inten
tional act was able to bring into existence its own object, we must also recognize that (1) and (2) have a parallel structure. The only difference is that while the act underlying (1) is directed towards a previously existing (act-independent) object, the act underlying (2) is directed towards an intentionally created (act-dependent) object.

The observation that certain kinds of intentional act are ontologically creative is of crucial importance for artefactualists. This is the observation on the basis of which they can argue that fictional characters have extremely minimal existence conditions (cf. Thomasson 2003).

According to their view, only the following two conditions must be satisfied for a particular character to come into existence in a fictional text $F$: (i) the author of $F$ must perform an ontologically creative intentional act that is directed towards a previously non-existent object, and (ii) the sentence corresponding to the creative act of the author must be included in the set of sentences composing $F$.

Our imaginary example seems to satisfy both of these conditions. In the foregoing discussion, we have supposed that (2) was the very first thought about Holmes in the history of literature. So we have conceived Doyle’s intentional act as genuinely creative. We have also supposed that (2) was a sentence token which occurred in a fictional text written by Doyle. Since
nothing more is required for character creation to succeed, we may conclude
that in performing (2) Doyle has created a new fictional character.

Although Holmes is thought in (2) as smoking a pipe, and in our real
world only persons are smokers, it would be folly to think that Doyle created
a person. Fictional characters are not denizens of the spatiotemporal world,
hence they are not persons. They have a somewhat peculiar ontological
status. Their existence, as we have seen, depends constitutively on worldly
processes and entities, like intentional acts and inscriptions, yet they are not
material in any sense. In this regard, characters belong to the same ontolog-
ical category as marriages, nations, institutions and other artefacts. All of
these sociocultural entities are contingently existing abstract objects. The
artefactual theory maintains, accordingly, that Doyle’s creative mental act
and the corresponding sentence token “Sherlock Holmes smokes a pipe” has
brought into existence an abstractum.

But this cannot be the entire explanation, since artefactualists are faced at
this point with an immediate problem. If Holmes is an abstract object, then
(2) can be paraphrased or translated, without loss of meaning, into (3):
(2) Sherlock Holmes smokes a pipe.
(3) The abstract object created by Conan Doyle smokes a pipe.

Of course, (3) is a nonsensical claim, one to which even the most committed
adherents of abstract objects could not assent. In attempting to solve this
problem, artefactualists may follow different strategies. First, they might
object that the paraphrase is inaccurate because the proper name “Sherlock
Holmes” and the definite description “the abstract object created by Conan
Doyle” have different modal profiles. As a rigid designator, the name
“Sherlock Holmes” refers to the same (abstract) object in all possible worlds.
Definite descriptions are, however, non-rigid designators, so in some nearby
worlds it is false that Holmes was created by Doyle. This is a well-known
objection of the direct reference theory against descriptive sentential para-
phrases and, as such, it has a standard answer. For example, the modal
difference between (2) and (3) may be eliminated by applying a suitably
rigidified version of the definite description. Inserting the actuality operator
@ or the sentence modifier actually in (3) enables the sentence to express
the same descriptive content in all of its uses in other possible worlds. Or, to
achieve the same effect, it might be claimed that the proper name ‘Sherlock
Holmes’ expresses a rigidified descriptive content which remains constant in
all of its possible tokenings. Artefactualists must therefore find another way
to solve the problem posed by the above paraphrase.

A better proposal is to say that, contrary to its surface appearance, (2) it is
not an assertion in the usual sense of the word. In fact, this is what Schiffer
(1996), Thomasson (2003, 2010) and other artefactualists contend. It seems
as if Doyle’s original sentence token attributed a property to an (abstract) object. But this is just a formal illusion, so to say. Authors of fictional texts do not make genuine assertions, they merely pretend to assert what they write. If it has been made clear that Doyle merely pretended to attribute a property to Holmes, then (2) can be paraphrased only in a way that makes explicit reference to this pretense. In applying the pretense strategy, the artefactualist can easily block the inference from (2) to (3) and say that the appropriate candidate for paraphrasing Doyle’s sentence is (4) or something similar to it:

(4) Conan Doyle created an abstract object and pretended to assert that it is a person who smokes a pipe.

The core idea, which most realist theories of fiction agree with, is to make a significant distinction between two kinds of assertion. Contemporary Meinongian theories provide perhaps the clearest example of how to motivate this distinction. On some neo-Meinongian views, fiction-internal assertions attribute nuclear or characterizing properties to fictional objects. Assertions within the story attribute such nuclear properties to Holmes as smoking a pipe or being talented. These are the properties which we associate directly with Holmes when we read Doyle’s story. Fiction-external assertions, in contrast, attribute extranuclear properties to objects. Among the extranuclear properties we find such properties as being fictional or being abstract. Thus, if we approach the story from an extra-fictional viewpoint, we may attribute to Holmes the property of being an abstract object. Other neo-Meinongians maintain that the distinction between fiction-internal and fiction-external assertions reflects the ways in which objects may possess their properties. While Holmes encodes the property of being a smoker, Günter Grass exemplifies that property. The term “encoding” is meant here to refer to the fact that Holmes has this property internally or as a constituent part of his (its) own being. The term “exemplifying,” on the other hand, is meant to indicate that Grass has the same property externally or non-constitutively.

If we now look back to (2), (3) and (4), we will understand the reason why artefactualists must subscribe to the fiction-internal/fiction-external distinction. They hold that fiction-internal assertions are capable to create abstract objects; yet at the same time they deny that these objects are literally ascribed any properties in the texts in which they occur. The only option which remains is thus to say that fiction-internal assertions merely pretend to make assertions. Nevertheless, when we literary theorists, linguists or philosophers talk about these abstract objects we do not talk about them in the scope of pretense. And this, in turn, presupposes that there is a pretense-free, external perspective from which we can make genuine assertions about the features of fiction-internal discourse. As Thomasson (2003) remarks, fiction-internal
occurrences of the proper name “Holmes” are non-referential, but these non-referential uses enable the name to refer to the character Holmes in external discourses.

These considerations lead, in the end, to an artefactual theory which admits that authors of fictional texts create literally fictional characters, but deny that the resulting texts contain literal assertions about these characters. This is, to say the least, a bit perplexing. The folk theory seems to be right in claiming that fictional texts do not have such contradictory features. In engaging with Doyle’s novel, it is a quite natural attitude to take sentences like (2) as telling us something about the personality and behavior of the character Holmes. Of course, we do not have to give up the artefactual theory in its entirety, simply because it conflicts with our ordinary ways of thinking about literary characters. This is especially so, if we realize that the source of the conflict lies in the conceptual apparatus of the theory. In the next two sections, I will argue that a slightly modified understanding of artefactualism can provide a non-contradictory view of our engagement with fictional texts and fictional characters without relying on the traditional fiction-internal/fiction-external distinction.

2. A Representationalist Alternative

At first sight, the act-object theory of intentionality seems to be the obvious choice to explain how authors can create new fictional characters. Thomasson (1999) thinks of intentional acts as potentially creative, so she can convincingly argue that certain kinds of authorial act bring into existence new intentional objects. The resulting art-ontological view holds, then, that fictional characters exist as dependent abstracta. As we have already indicated earlier, the creation of Holmes is thought of as depending both on the ontologically productive mental act performed by Doyle and the sentence token which corresponds to Doyle’s act.

The second part of the latter claim, however, rests on a questionable presupposition concerning the semantics of fictional sentences. In order to understand correctly what this presupposition amounts to, let us consider once again what the artefactualist has to say about the ontological relation between characters and the sentence tokens on which their coming into existence and their continued existence is (partly) dependent. Given their realist ontological commitments and their accompanying ideas about contingently existing abstract objects, artefactualists have to distinguish characters from literary texts in which their names are (first) mentioned. Although their existence conditions include facts about linguistic objects, fictional characters are supposed to be abstracta which are not linguistic by their nature. But, then, what kind of abstract object are they? Voltolini (2006) identifies them
with compound abstract entities. On Voltolini’s view, characters are constituted by an instantiated make-believe process-type and a fiction-internally ascribed set of properties. Thomasson who also denies the linguistic nature of characters, proposes another possible account, namely that they are ontologically similar to marriages, laws and other abstract cultural artefacts (cf. Thomasson 1999, 2003). From a Thomassonian point of view, the linguistic object – the token name “Holmes” – used by Doyle for expressing his creative thought and the resulting Holmes-artefact are related but distinct objects. They are related because they are dependent on exactly the same mental basis, namely on Doyle’s intentional acts. Yet they are distinct, because they are instances of different kinds of abstractum.

Unfortunately, these conceptions presuppose rather than explain the ontological status of characters. The core problem may be stated in the following way. The semantic content of Doyle’s sentence token and the Holmes-artefact are arguably different kinds of abstract object. If the Holmes-artefact is (partly) dependent on but not identical with the semantic content of the sentence token (2), then there must be a causal or inferential relation actually connecting them. But so far neither Voltolini nor Thomasson have attempted to explain how this relation could be defined in an artefactualist framework. Moreover, the available linguistic data do not seem to support such an explanation.\(^8\)

I think the simplest way to get rid of this problem is to hold that Sherlock Holmes is nothing more than a linguistic construct. Let us say that the sentence token “Sherlock Holmes smokes a pipe” represents Holmes as being a certain way. Doyle’s sentence represents Holmes as smoking a pipe, so we may say that it offers us a particular Holmes-representation. Note that the term “representation” is used here in a specific, non-relational way. There is nothing new in this usage. In recent years, various attempts have been made in the relevant philosophical and linguistic literature to develop a non-relational interpretation of representation.\(^9\) For example, Burge (2010) points out that representation can transpire even in cases where there is nothing that could be represented. In Burge’s jargon, representing is always representing-as-of. A predicate or a proper name \(X\) has this underlying structure independently of whether there is a candidate for being a representatum of \(X\). Sainsbury (2012) is of the same opinion when he remarks that the claim “\(X\) represents \(Y\)” does not imply that there is something, \(Y\), such that \(X\) represents it. In this sense, representing Holmes is not equivalent with expressing a singular proposition containing Holmes as a constituent. A sentence token can represent Holmes as a smoker even if there is no such person as Holmes. Moreover, the Holmes-representation is interpretable without any reference to language-independent objects, be they concrete or abstract. It is tantamount to say that fictional characters come into being as self-standing, non-relational linguistic representations.

74
Now we can take a further step in this direction. Doyle’s novel represents Holmes in a great variety of ways – as being a smoker, as being a detective, as living on Baker Street, and so forth. Thus, if Doyle is the author of the fictional text $F$, then Holmes can be identified in $F$ with the elements of the property set $P$, where $P = \{ p_1, p_2, \ldots, p_n \}$, and $p_1, p_2, \ldots, p_n$ are the properties explicitly represented by the sentences of $F$.

According to $F$, Holmes is a smoker. We know, however, that smokers often have lungs that are gross and black. So Holmes may also have gross and black lungs. But $F$ is absolutely silent about this matter. $F$ represents him neither as having gross and black lungs, nor as having healthy lungs. This is not to say that Holmes is an incomplete representation, it means simply that he is not explicitly represented either way. But some elements of the property set $P$ may jointly imply that he has healthy lungs. Then, it would be certainly correct to state that Holmes is represented implicitly as having healthy lungs. Compare this with the case when someone claims that Holmes is a computer program. We would definitively reject this claim by saying that Holmes is represented neither explicitly nor implicitly as a computer program in $F$.

In order to decide such cases we may introduce the property set $P^*$, where $P^* = \{ p^*_1, p^*_2, \ldots, p^*_n \}$, and $p^*_1, p^*_2, \ldots, p^*_n$ are the implicit properties derived from $P$ through an appropriately designed interpretative method $M$. Since it is not a trivial task to develop such a method as $M$, we must be content to assume that it can in principle be developed. Under this assumption, the Holmes-representation of $F$ may be seen as completely determined by the property sets $P$ and $P^*$. Or, in other words, thinking about Holmes means nothing else than mobilizing a non-relational representational complex combined from certain elements of $P$ and $P^*$. This does not mean that Holmes is a set-correlate of the union of $P$ and $P^*$, since we cannot ascribe properties to fictional characters: characters cannot be subsumed under the traditional ontological category of language-independent objects. The central claim is, rather, that Holmes exists in $F$ as a maximal representational combination of the elements of $P$ and $P^*$.

It is important to keep in mind that the terms “property” and “property set” are not used here in their neo-Meinongian sense. We can say, if we like, that in writing his novel Doyle ascribes a particular property (say, being a smoker) to his protagonist. But this idiom is to some extent misleading, since in the context of $F$, property-ascribing predicates like “is a smoker” can acquire only non-relational interpretations. There is no language-independent object (Holmes) to which the text may ascribe a language-independent property (being a smoker). Thus, in contrast to the neo-Meinongian theory which conceives property attribution as a relational act, the present view
acknowledges properties only at the level of non-relational linguistic representation.

At this point, the following objection may be raised. If Holmes is a linguistic representation created by Doyle, then we are back where we started. The problem now seems to be that we are forced to accept (5) as a correct paraphrase of (2):

(5) The Holmes representation created by Conan Doyle smokes a pipe.

Undoubtedly, (5) sounds as bad as (3). But (5) is misleading in that it makes it seem as if the property of smoking a pipe were a property of the Holmes representation. This cannot be the correct interpretation. Being a member of $P$, the property of smoking a pipe is in fact a constitutive part of the Holmes representation of $F$. Hence it is incorrect to think that Doyle’s creative act ascribed a property to an abstract linguistic construct. We can avoid this misleading interpretation by focusing on the physical basis of the Holmes-representation:

(6) The sentence token created by Conan Doyle represents Holmes as smoking a pipe.

(6) reveals two things about the circumstances under which the fictional character Holmes was created. First, Doyle’s intentional act was ontologically productive because it created a particular mental representation. One may conjecture that this mental representation was the original content of his creative act. This content can be seen, in turn, as the way in which he created that mental representation. The “way” of the creative act can be described adequately by an adverbial phrase: Sherlock Holmes was represented in Doyle’s thought in a smoking-a-pipe-wise. Or, more correctly, we can say that Doyle performed a mental act in a Sherlock-Holmes-smokes-a-pipe-wise. This particular way of thinking was what individuated the content of the creative act in question. Second, in order to successfully expressing his thought in a publicly accessible form, Doyle must have written down the sentence “Sherlock Holmes smokes a pipe.” Hence it is not unreasonable to think that this token sentence acquires its content from the mental act that immediately precedes it. And given that the mental act represents Holmes in a non-relational manner, the token sentence must represent Holmes in a non-relational manner, too. The result of the writing act can thus be regarded as the physical endpoint of character creation. There is no need for supplementing this account with a further inferential step: (2), just as it is in itself, is able to bring into being Holmes as a publicly accessible non-relational representation.
3. In Closing

It thus seems that there is little reason to reject the artefactual theory of fictional characters. Paradigmatic examples of fictional creatures like Sherlock Holmes, Fyodor Karamazov or Madame Bovary are contingently existing abstract objects. Artefactualists are fundamentally right in their emphasis on the created nature of these entities. But I think we remain closer to the folk view if, contrary to standard artefactualism, we regard fictional characters as linguistic representations.

One obvious advantage of this revised version of the artefactual theory is that the terminological distinction between fiction-internal and fiction-external assertions becomes superfluous. Consider first the easy case in which someone tries to understand what a fiction-internal sentence asserts. When we read the sentence “Holmes is a detective,” we come to know that Holmes is represented in the text as being a detective. This is a natural-sounding interpretation. A much harder case is when we read the apparently fiction-external sentence “Holmes is a fictional character.” Does the interpretation of this sentence actually involve an external perspective? It is not necessarily so. It just needs to be recalled what the term “fictional character” means in the present context. According to the present account, fictional characters are linguistic representations. And the term “representation” is intended to mean, in turn, “to be represented in a certain way.” Thus, the sentence “Holmes is a fictional character” informs us that Holmes is represented in the text as being a certain way. It does not provide a particular representation; rather, it reminds us that Doyle’s character has a representational nature. Since this seems to be in agreement with the former result, one may conclude that Holmes is seen through the same interpretative perspective in both cases.

A further and related advantage of the present proposal is that it does not entail the controversial claim according to which authors of fictional texts merely pretend to make assertions. The pretense approach to authorial assertions is used by Thomasson and other artefactualists as a last resort. Its only role is to assure us that fiction-internal assertions do not concern abstracta. But once one thinks of the fiction-internal/fiction-external distinction as theoretically insignificant, even this limited role seems likely to evaporate altogether. Thomasson and her followers are right in saying that authorial assertions do not ascribe properties to abstract artefacts. From a representationalist point of view, however, one can provide a complete explanation for this phenomenon without involving the notion of authorial pretense.

The present proposal requires only a light revision of the standard artefactualist framework. It leaves untouched the abstract status of fictional characters, but makes an attempt to state more precisely what kind of abstract objects they are. The core idea is that fictive individuals in literary
artworks like Sherlock Holmes, Fyodor Karamazov and Madame Bovary are to be thought as non-relational linguistic representations. Though I am fairly well convinced that this idea allows a simpler and more uniform theoretical treatment of fictional discourse, some important details may be lacking. Arguably further analysis is required to account for the naturalness of the representationalist view: is it indeed the case that everyday consumers of literary works think of fictional characters in terms of representation? But this will be the subject of another paper.

NOTES

1. The most elaborated and consistent version of the artefactual theory is to be found in Thomasson (1999). In responding to her critics, Thomasson has revised some of her original ideas over the years (cf., Thomasson 2003, 2010). In what follows, I shall focus primarily on her view in my discussion.

2. In fact, the proper name Sherlock Holmes occurs for the first time in the novel *A Study in Scarlet* (1887/2004). The first sentence which contains the name is to be found on page 5: “You don’t know Sherlock Holmes yet.”

3. For similar observations, see Yagisawa (2001) and Brock (2010). The charge of circularity is not completely unfounded, but this is irrelevant for the present purposes. I shall argue for a version of artefactualism that avoids this problem entirely.


5. It is worth to note that artefactualists borrow the idea of pretense from anti-realist theories which deny the existence of fictional characters. This is not a problem in itself. But it is hard to understand how can Doyle intend to create literally an object if he merely pretends to assert something. Wouldn’t it be more appropriate to claim that he pretends to create an object? I think it is a legitimate question, but to answer it would require a separate study.

6. On the differences between Neo-Meinongian theories see, for example, Zalta (1992).

7. Anti-realist theories of fictional characters must also adopt some variant of this distinction, because they have to explain how we can talk about fictional characters in spite of the fact that they do not exist. On this theme, see Friend (2007).

8. According to Kertész and Rákosi (2012), linguistic data can be considered to be more or less reliable “truth-candidates.” Data are supported by their sources only to a certain extent, hence they are not true with certainty. And similarly, theoretical claims is linguistics are usually not true with certainty, but more or less plausible. If we follow this account, the artefactualist’s claim can be regarded as having a low plausibility value.


10. Following Kriegel (2008), we may call this account *content adverbialism*. On this account, thoughts about fictionalia are related to representational contents, but, and this is the main point, representational contents in themselves are not relational.
11. One of the central issues which must be addressed by artefactualists concerns the problem of name-introduction procedures. Adherents of the artefactual theory typically think that a certain version of the baptism-based theory of ordinary proper names is able to provide an account also for fictional names. Of course, the baptism-based theory cannot be straightforwardly adopted to fictional cases because of the absence of objects which could be baptized. One possible solution is to follow Thomasson (1999), who argues that characters acquire their names in a quasi-indexical way. For more details of this topic, see, among others, Braun (2005) and Lamarque (2010).

REFERENCES