NORTHWESTERN ITALIAN
PHILOSOPHY CONSORTIUM

XXX Cycle

Curriculum : Mind, Science and Language

MENTAL FILES
AND FICTIONAL OBJECTS

PhD Thesis

Supervisor:

Prof. Carola Barbero

Candidate:

Enrico Grosso

November 2018

CONTENTS

Chapter 3 : Fiction and Mental Files

INTRODUCTION

Philosophers have been interested in the problem of fictional objects from the very beginning of analytic philosophy. The debate that opposed Russell to Meinong is well known: if the former refused any ontological commitment to such objects, the latter was among the first ones to explore the metaphysical side of the issue. We should not forget another important contribution. With his reflection on the notion of sense, Frege opens up a further field of inquiry that relies on a question of cognitive nature: what do we think when we think about fictional objects? In recent years, for matters that have nothing or little to do with fiction, the Fregean theses have returned to the forefront of discussion. His two-dimensional semantics has led to the development of the mental files theory. I shall argue that this theory can open some interesting perspectives that enable us to investigate better the problems that fictional objects carry with them. This work aims to provide an analysis of the cognitive tools that we use to represent literary characters in our mind: only after solving this preliminary task we can give adequate answers to the other questions.

The first chapter is devoted to the presentation of the main philological theories that have been proposed over the last century. They can be clustered into two groups, depending on whether they claim that entities such as fictional objects exist or do not. As far as the antirealist approach is concerned, the discussion has mainly focused on semantics. Fictional antirealists have drawn inspiration from the Russellian idea of paraphrase, by proposing more and more sophisticated treatments of sentences apparently involved with fictional objects in order to account for their intuitive truth-values. However, neither the use of intensional operators, nor the appeal to the notion of make-believe games, succeed in giving a fully satisfactory account of all cases, especially those related to external metafictional sentences, i.e. sentences uttered with a non-conniving use with the purpose to express genuine truths or falsehoods about fictional objects, regardless the story they belong to. On the other side, fictional realists acknowledge that names of purely literary characters have reference, but they must explain the nonexistence datum, i.e., the fact that fictional entities fail to possess a

spatiotemporal location in our world. Possibilist theories accommodate the actual nonexistence of fictional objects by identifying them as a sub-class of possible objects. In some possible worlds, literary characters exist as concrete individuals that instantiate what is said about them in the relevant narrations. Conversely, neo-Meinongian philosophers conceive fictional objects either as concrete, albeit nonexisting, correlates of sets of properties, or as eternal and unchangeable abstract objects along the model of Platonic attributes. Finally, artifactualist theories hold that fictional entities are abstract cultural artifacts that originate from a creative act of imagination: literary characters are bound idealities, in the sense that they depend either historically rigidly from their authors and constantly generically on literary works and other media that feature them. It is made clear that, even if each theoretical proposal contributes to sharpen and enrich the discussion on this topic, no one is exempt from difficulties and succeeds in prevailing over the alternative accounts. The current debate seems to have reached a deadlock from which it is hard to get out. For this reason, I suggest that it is time to change the perspective and, following the direction pointed out by Frege, tackle the problem of fictional objects by starting from the cognitive side.

In the second chapter, I examine the diatribe between descriptivism, i.e. the view that our mental relation to individual objects goes through the properties of those objects, and singularism, i.e. the view that our thought is about individual objects as much as it is about properties, with the purpose of framing the theoretical background that gives rise to the theory of mental files and to clarify the reasons that led to the recovery of the Fregean notion of sense, intended as a non-descriptive mode of presentation. The term “mental file” was originally introduced by Perry in the eighties and it has been widely used since then. I shall base my analysis on Recanati's important contributions in this field. According to him, a mental file is a cognitive structure that enables us to create a mental representation of what we take, internally, to be a singular object of the external world. The primary and basic function of the file is to store information (or misinformation) about that object, information that the subject acquires in a non-descriptive way by means of an epistemically rewarding relation (that is, a relation of acquaintance) that she entertains with the object. Acquaintance can be direct, when we stay in a perceptual contact with the object, or mediated by the existence of a communicative chain, as it happens for people who lived in the past, or by contextual relations, as in the case of indexicals. It is then showed how mental files are applied to solve traditional problems, like those involving identity statements and belief ascriptions, as well as

other well-known Kripke's puzzles, afflicting either descriptivism *à la* Russell and singularism *à la* Mill.

In the third chapter, I explain how Recanati's theory can be applied to the field of fictional objects. The starting point of my analysis is the assumption that in fictional stories we can find both empty and non-empty singular terms, i.e., in addition to names that are entirely made up, fictions deploy names of individuals that exist in the actual world. Following Friend, I assume that a) when real names are used in fictional contexts, they maintain their ordinary reference and b) we get fictional information about the referent of that names. I argue that, by means of the theory of mental file, we can provide a unitary cognitive account of how names work in fiction. However, unlike Recanati and Friend, I claim that the crucial notion we need is not the one of regular file, presented in the second chapter, but the notion of vicarious file, i.e., a file that has a meta-representational function which allows us to represent another subject's way of thinking about an object. This kind of file is characterized by an indexed structure, $<f, S_2>$, where the index refers to the other subject whose own file the indexed file stands for or simulates. Vicarious files are needed to store all and solely those pieces of information that we associate to the world of the fictional story, as participants of a certain game of make-believe, without combining them with our knowledge of the actual world that is, instead, contained in our regular files. When we read a novel displaying the name of a real person, as "Napoleon" in Tolstoy's *War and Peace*, we acquire (fictional) information about that individual, by generating a vicarious file indexed to the story. The vicarious file is vertically linked with our regular file on Napoleon, but their content is not joined: we keep separate fictional and real information about the referent and we can exploit both files, depending on the situation. Consequently, I claim that files still work in the same way when we deal with names of purely fictional characters: we open a vicarious file and we begin to store information in it. However, in this case, it is a free-wheeling, or unloaded, use of the vicarious file, since we lack a regular, acquaintance-dependent file in our mind to be linked to. Thus, no reference to objects of the actual world is made and no parallel between fiction and reality is possible, since the only information we have is the one derived from the story.

By means of this conceptual apparatus, I can accommodate intuitions about fictional characters that, at first glance, seem hard to keep together, or even explicitly contradictory. On the one hand, the identity of a fictional character is linked to the particular work of fiction in

which it appears: Ulysses is described in one way in Homer's epic poems, in another way in Virgil's *Aeneid*, and in a still different way in Dante's *Divina Comedia*. Each version of this literary character is distinct from the others and has specific properties. On the other hand, it is natural to think of Ulysses as a unique fictional character that moves through different works, even though in the migration it changes, even significantly, its features. My thesis is that vicarious files gather into networks according to the causal relations that subsist between literary works. For the identity of literary characters at the metafictional level, i.e. in the comparison between two or more works of fiction, I provide a criterion to determine whether a character is new or represents a further version of an already existing character, a criterion that enables us to address even phenomena of fusion and fission of literary characters. As a final consideration, I focus on regular files in order to account for the difference between fictional and metafictional sentences or, to be more precise, between conniving and non-conniving uses of fictional sentences. I argue that, from a cognitive point of view, the distinction between the two types of sentences can be explained in terms of the mental files that are engaged. We use vicarious files when we talk about a character from a perspective that is internal to the fiction. On the contrary, we use regular files when we want to say something true or false about the very same character regardless of any specific practice of pretense or make-believe. However, regular files refer to fictional entities not as flesh-and-blood individuals, like they are usually depicted in their relevant stories, but as abstract objects whose nature, whatever it is, still need to be fully determined.

CHAPTER 1

THE PROBLEM OF FICTIONAL OBJECTS

1.1 The problem of the nonexistent

Suppose to go into a library. Not a library specialized in scientific texts, like those that you can find at the University. Just think about an ordinary neighborhood library. The first thing you will notice is the impressive number of books labeled as novels or works of fiction. These books come from every historical age and all over the world. And more surprisingly, in such neighborhood library, the number of works of fiction probably exceeds the number of any other kind of books, such as scientific, philosophical or historical essays. Let us consider another example. According to the Italian Publishers Association (AIE), 66.757 new books were published in Italy in 2017. Works of fiction are almost a third of them, 19.860 to be precise. The situation is likely to be similar in other countries. This is a fascinating phenomenon, for philosophers generally believe that we use language to talk about the reality, i.e., to correctly or incorrectly describe states of affairs of the world. For example, I can claim that Torino is an Italian city. In this case, I am making a true statement. Or I can maintain that Torino is the biggest city in Italy. In this second case, what I am saying is false. However, in both cases it is undeniable that I am talking about a particular object of the world and I am ascribing to it (rightly or wrongly) a certain property. Indeed, the name "Torino" refers to a real city, Torino, which has some properties, as "being an Italian city" and not some other properties, as "being the biggest city in Italy". If I have doubts, any geography manual will provide all the information I need. But when I read in *One Hundred Years of Solitude* that events take place in Macondo, I cannot find this city in a geography manual, for Macondo

does not exist anywhere in the world. Márquez's book is not reporting some actual state of affairs. So, what is the author talking about when he talks about Macondo? And then, is the author saying something that is true or false?

Philosophers have to address the problem that language is not always used to speak about real things, and that indeed a considerable amount of discourses and texts do not. As Tim Crane underline, our talk of the nonexistent is pervasive and ubiquitous:

> Using names for non-existent people and things is rooted in our talk and thought about the world. This is partly because of our interest in fictions – stories which are (often) precisely about things that do not exist. But it is also partially because of the fragility of our epistemic endeavours. We create theories about the world and hypothesize that there are all sorts of things. Sometimes we are right, sometimes we are wrong. But when we are wrong, our words often work in similar ways, they have similar roles in our languages, and we still are able to talk about these things. (Crane 2013: 14)

The puzzle of the nonexistent does not concern only literature, but also other domains of discourse, as science, that apparently should deal only with things of the real world. For this reason, we can distinguish between two ways we can talk and think about the nonexistent[1]:

> (i) Error: this is when we talk about things which have been genuinely supposed to exist, but do not. Le Verrier's supposed planet Vulcan, phlogiston, the fountain of youth, and similar things have all been in this category at one time to another, as are the objects mistakenly thought to exist by those who are hallucinating, whatever the cause.
>
> (ii) Fiction: this is when we think about things which we know do not exist. Many characters and places in novels, plays, and movies do not exist; many objects of fantasy and imagination, and so on, likewise – they all belong here.
>
> (Crane 2013: 15)

The first category typically includes all those entities postulated by scientists that proved to be false: the planet Vulcan, hypothesized by Le Verrier on the basis of his observations of the orbit of Mercury; the fire-like element called phlogiston, which was supposed to explain the processes of oxidation and combustion, subsequently denied and abandoned after the

[1] The original distinction is proposed by Donnellan (1974: 5).

publication of the mass conservation law by Antoine Lavoisier; epicycles, used by Ptolemaic astronomers to explain the variations in speed and direction of the apparent retrograde motion of the Moon and planets; and so on. Instead, what characterizes the second category, compared to the first one, is that fictional discourses are uttered with the awareness that the objects we are talking about do not exist. Clearly, members of the second category are literary characters, called by philosophers either *fictional entities* or *fictional objects* or *ficta*. This class includes not only animate beings, like people (Sherlock Holmes, Anna Karenina, Romeo and Juliet), animals (Moby Dick, Puss in Boots, Ulysses' faithful dog Argos), or monsters and chimeras of various kinds (Pegasus, Polyphemus, centaurs, sirens, ghosts, fairies and elves), but also inanimate object and fictional places (the philosopher's stone, the Holy Grail, Excalibur, the island of Utopia, Pinocchio's Land of Toys, and so on). Note that, even if the word "fiction" has been used, the class does not include only those literary creations that have been introduced with the aim to delight or entertain. Nonexistent objects of various kinds also appear in serious philosophical and scientific discourses. An interesting field of discussion, for instance, is the one concerning the role of fiction in science. According to some philosophers, elements of fiction are employed as epistemic tools to provide a better representation of the world and to construct idealized predictive models[2]. Some examples taken from physics are points mass (defined as centers without spatial extension and with no other properties except that of having mass), force fields (a vector field that describes a non-contact force acting on a particle at various positions in space), electromagnetic waves (synchronized oscillations of electric and magnetic fields that propagate through space-time at the speed of light) and many others. This is a vast and complex debate, largely independent of the purposes of this work, that I shall leave aside in order to focus on the more traditional examples of fictional objects[3].

The categories of errors and fictions should not be considered as hard and fast sets, since there are elements that, at first originated by mistake, find their fortune as a source of inspiration for literary works. Vulcan, for instance, although it is now a forgotten scientific hypothesis, continues to appear in works of science fiction, like the Tv series *Star Trek*.

[2] This debate in the philosophy of science originates at the beginning of the twentieth century with Vaihinger's work (1924). For a greater insight and recent contributions on the debate, see Sainsbury (2009) and Teller (2009).

[3] Vaihinger himself (1924) distinguishes scientific fictions from mythological or aesthetic fictions. Whereas scientific fictions are virtuous, in the sense that they are "legitimatized" errors which have an epistemic justification, non-scientific fictions, called *figments*, are made up with the aim to entertain, delight, or evoke some other kind of emotion.

Examples even more emblematic are provided by myths and folktales: Greek gods and heroes, like Zeus or Achilles, once believed to be real, are now regarded as literary and cultural creations. Arguably, the reverse shift can also take place. I suggest one example: the memorable Mercury Theatre's broadcast, aired on October 30, 1938, during which the young Orson Welles interpreted a radio adaptation of *The War of the Worlds*, by H. G. Wells. The broadcast was presented as a series of simulated news bulletins and the program triggered panic in most of the United States, for many radio listeners believed that Martians invaded the Earth. We can say that, in such a case, even if Martians have been originally thought as merely fictional, people mistakenly supposed them to really exist and the nation to truly be under attack. All these examples illustrate that things may belong to both categories across time.

So far, I have circumscribed the field of investigations that the present work aims to investigate. Before going on, I wish to point out that the two categories we have presented are not exhaustive of the whole area of the nonexistent since there are cases that cannot straightforwardly fall into neither the class of errors or the class of fictions. Paradoxically, the domain of the nonexistent is much wider than it may appear at first sight. It includes, or it may include, depending on the philosophical perspectives one may sympathize with, diverse arrays of objects that have no spatiotemporal existence in our world. For instance, (im)possible objects (*im-possibilia*), which represent one of the most debated topic analytic philosophy[4]. Some classic examples are the golden mountain, proposed by Meinong as a case of object that does not actually exist but is metaphysically possible, or the round square, proposed as a case of metaphysically impossible nonexistent object. Abstract objects constitute another controversial field that has traditionally attracted the attention of philosophers, at least starting from the fundamental contributions by Quine (1948). Within this dispute, a special place is occupied by the discussion around the nature and the function of numbers and other mathematical objects. For some philosophers, mathematics has a mere epistemic function, it is a useful tool to investigate the world like other fictional constructions we can find in scientific theories: “Mathematics, as a whole, constitutes the classical instance

[4] Of course, the notion of possible objects becomes interesting only if we do not identify the category of possibility with the category of actuality: “There is a wide-spread conservative view on objects, which says that any object is an actual object. In other words, the adjective ‘actual’ is redundant, for it excludes no object. From this it follows that non-actual possible objects are not objects, that is, they are nothing. Thus on this view, the adjective ‘possible’ is equivalent to ‘actual’. [...] This makes the notion of a possible object, or equivalently the notion of an actual object, uninteresting. The notion of an object is the basic notion and does all the work.” Yagisawa (2018).

of an ingenious instrument, of a mental expedient for facilitating the operation of thought" (Vaihinger 1924: 57). Mathematical knowledge, thus, does not carry any ontological commitment. Other philosophers claim that existence is not limited to existence in space and time, so we can make room in our ontology for numbers and other abstract objects[5]. A third class of objects whose nature is problematic is provided by those things that are, in one way or another, connected with the time. Entities that once existed, but that exist no longer, like people who died, cities that were destroyed, animals that became extinct. Entities that do not exist in the present time, but that will exist, perhaps, in the future, like babies that will be born in 2020 or the first self-driving car that will be put on the market. In our everyday conversations, we usually talk about the past or the future, although both are metaphysically uncertain: for some philosophers, existence is limited to the present time, so there are not past or future things[6]. It is not possible to address in detail all these subjects that I briefly mentioned here, for an in-depth analysis of them would bring too far from the scope of this work.

One last thing must be added. Among the cases that can be put neither in the class of errors nor the class of fictions, there are ones to which philosophers have so far paid little attention. Consider the following example: the Donation of Constantine is a forged Roman decree granting the Catholic Church ownership of vast territories within the western Roman Empire. The document stated that emperor Constantine made this generous gift out of gratitude to Pope Sylvester I, who had converted him to Christianity and had cured him of leprosy. For centuries, the Donation was used by the Church to legitimated its domains in Italy. Only in the XV century, the Italian philologist Lorenzo Valla unequivocally demonstrated that the Donation was a fake. Unlike Orson Welles' broadcast on the Martian invasion, which was performed as a work of fiction but taken as true due to a misunderstanding, the Donation of Constantine is a fictional object intentionally made up to deceive and lead to suppose that it genuinely exists. Hoaxes, scams, frauds and the now sadly known fake news are all problematic cases, since their ambiguous nature challenge any sharp distinction between fictions and errors. I leave these cases for future investigations.

---

[5] For an introduction to the debate, see Rosen (2017) and Horsten (2018).

[6] For an introduction to the debate see Ingram & Tallant (2018). The debate between Presentism (the view that only present things exist) and Eternalism (the view that there are such things as merely past and merely future entities) is seen as strictly connected to the debate between Necessitism (the view that necessarily everything is necessarily something) and *Contingentism* (the view that denies that necessarily everything is necessarily something). At this regard, see Williamson (2013).

1.2 Historical background

There are two basic questions we can ask about fictional objects, as Voltolini put it:

> [...] the question about what is an object of fiction is a *metaphysical* question, at least if by "metaphysics" we mean a question about the *nature* of things, or rather about what kind of things are things of a certain kind. [...] The metaphysical question can be separated from the *ontological* question, i.e., from the question of whether *there are* things of a certain type, whether the general catalog of what there is also contains things of this kind. As far as fictional entities are concerned, the distinction is appropriate since the real existence of such entities is a very controversial issue. (Voltolini 2010: 52, *my translation*)

Following Thomasson (1999: 5), the metaphysical question amounts to ask: what would fictional entities be, if there were any? Such a question presupposes the ontological one. For it makes sense to wonder about the nature of fictional entities only if we have preliminarily replied that yes, these things are part of the overall domain of what there is. However, regardless the approach one may prefer, all theories must account for the so-called *nonexistence datum*, i.e., the fact that *ficta* do not exist in the world as other objects, like tables, houses, people, dogs, cats, etc., do exist:

> To [the metaphysical] question different answers have been proposed. But however much they differ, they all try to accommodate what seems to be an intuitive datum facing philosophers who theorize about fictional entities: these entities lack existence, or at least existence as ordinary physical objects. According to this datum – call it *the nonexistence datum* – paradigmatic objects of fiction like Hamlet and Holmes do not exist. We also appeal to nonexistence in this sense when we want to dispute the view that some alleged individual is a genuine historical figure, thereby underlining our view that a search for such entities would be in vain: we might say, for example, that King Arthur does not exist. (Voltolini & Kroon 2016)

From a semantic point of view, the question could be put in this way. There are expressions, the so-called *singular terms*, which are generally used to refer to single objects of the world[7]. Proper names of people, animals, cities, places, etc., are paradigmatic examples of singular terms. Thus, the name "Socrates" refers to a philosopher who lived in ancient Athens and died after drinking hemlock. The name "Paris" refers to a city, which is the capital of France. The name "Mount Everest" refers to the highest mountain on Earth. However, not all proper names have, *prima facie*, a reference. Typically, these are the names of literary characters, like "Ulysses", "Sherlock Holmes", "Anna Karenina". We can call them fictional names. Another well-known class of expressions that are generally used to denote single objects, but that can fail the task, are definite descriptions like "the current king of France", "the winged horse", "the Holy Grail", "the round square". Nonetheless, all these non-referring singular terms are not considered, by the overwhelming majority of philosophers, as completely meaningless expressions, like "andkanfskod" or "abrioc". For they can compare into perfectly understandable sentences that also appear to be true or false. The nonexistence datum is indeed given by the intuitive truth of negative existentials, like:

1) Sherlock Holmes does not exist

Philosophers that do not believe that there are any fictional entities (*fictional antirealists*) provide a mere ontological reading of negative existentials: saying that fictional entities do not exist amounts to say that there are not such kind of things. On the contrary, philosophers that believe that the overall ontological domain also includes fictional entities (*fictional realists*) provide a metaphysical reading of the datum, namely that fictional entities have the property of not existing. Nonexistence is then understood in different ways.

The current discussion around these problems begins with the very origin of analytic philosophy[8]: between the late nineteenth and early twentieth centuries, Frege, Russell and Meinong gave some fundamental contributions to the debate, and anyone who came after them had to deal with arguments and solutions that these three philosophies have proposed.

[7] For a more detailed discussion on singular terms, see chapter 2.

[8] Of course, antecedents are not missing and we could go back in the historical reconstruction up to the ancient Greek philosophy. See Eklund (2017).

Frege endorsed a two-dimensional semantics that splits the intuitive notion of meaning into two distinct components: a semantic component, the referent, and a cognitive component, the sense. Thanks to this distinction, he can acknowledge singular terms that have sense, but not reference. In fact, the sense determines the reference, but not *vice versa*: whereas any referring expression, like "Socrates" or "the oldest man in the world", must have sense, fictional names like "Ulysses" and definite descriptions like "the current king of France", have sense without having reference. In other words, we can say that sense is a necessary condition for reference, but not a sufficient condition, too. Meinong, instead, took the opposite direction. For him, in language there are not singular terms without reference, for every expression has one, albeit the referent can be a nonexistent object. Finally, Russell disagrees with both: he tries to give up either Frege's senses and Meinong's nonexistent objects, by equating names to definite descriptions and treating the latter, from a logical point of view, as "incomplete symbols" that introduce a quantificational structure into sentences in which they occur.

As we have said, these three positions have deeply influenced the debate about the problem of the nonexistent. One reason for their fortune is that they somehow mapped the fundamental options available to philosophers when they explore this field of investigation. In fact, while Frege and Russell agree in defending an antirealist perspective, Russell and Meinong both refuse the Fregean distinction between sense and reference in favor of a monostratal semantics[9]. However, the contraposition between Meinong and Russell had much greater historical weight, partly because the ontological dispute has been seen as more relevant, partly because Frege talked about fiction only marginally in his writings, whereas Meinong and Russell faced each other harshly on this topic. Fictional antirealists have drawn inspiration from the Russellian idea of paraphrase and have proposed different ways to analyze sentences containing non-referring expressions in order to account for their intuitive truth-values. On the other side, fictional realists have more or less directly included elements from Meinong's thesis in their theories, by conceiving fictional entities, from time to time, as nonexistent objects, abstract objects or artifacts. The only relevant exception is represented by the possibilist approach, which will be taken into consideration as well. In the rest of this chapter, we will present the ontological and metaphysical debate in its various ramifications[10].

[9] For an historical reconstruction of the thesis hold by Frege, Russell and Meinong, see Coffa (1991: ch. 4-5) and Orilia (2002: ch. 2, 5-8).

1.3 Antirealist perspectives on fictional objects

1.3.1 The Russellian road

Since the metaphysical problem presupposes the ontological problem, we will start our survey with an overview of the main antirealistic approaches. In fact, the claim that fictional entities do exist is far from being universally accepted. A large number of philosophers hold that we can account the various intuitions we have about them without committing ourselves to entities whose nature is so much odd and controversial. We are going now to see the arguments that have been proposed in support of this view.

Most of the debate is focused on the semantic analysis, since the basic idea that there are such things as fictional entities originates from our ordinary way to speak about them[11]. As Crane observes:

> So what is realism about fictional characters? Trivially, it is the view that fictional characters are real. Someone might hold this view because they want to explain the fact that fictions hold certain things to be true of fictional worlds or characters and certain things are false: that, for example, it is true that Sherlock Holmes is a detective, that Siegfried killed the dragon, that Macbeth was ambitious, and so on. After all, it is not true in any sense that Holmes is an aerobic instructor, or that the dragon killed Siegfried. A correct account of the content of a fiction will show that some of the claims within the fiction are true of the world it talks about. This kind of realism might be based on the simple view of truth and predication. "Sherlock Holmes is a detective" is true because the object referred to by the name "Sherlock Holmes" has the property referred to by the predicate "is a detective". (Crane 2013: 76)

[10] For the discussion that will follow, I am in debt with the lectures on *The Philosophy of Fiction*, held by Kroon, Voltolini, Barbero at the University of Torino in 2015.

[11] However, ordinary language also grounds the very opposite intuition, as we have seen for negative existentials.

Realists can easily account for the truth-conditions of these sentences, since they maintain that names of fictional characters have reference. On the contrary, antirealists claim that names like “Sherlock Holmes” and “Siegfried” do not refer to anything: they are *empty names*. Then, if I utter:

2) Sherlock Holmes is a detective

such a sentence cannot concern a certain real individual, Sherlock Holmes, and ascribe to him the property of “being a detective”; for there is no such individual. So, it seems, (2) has no truth-conditions: it is meaningless or, at most, it expresses a “gappy” or “incomplete proposition”[12]. However, it is not plausible to claim that all sentences containing fictional names are meaningless, or even false, as we will see later. Indeed, the challenge for antirealists is to show that, at least, some of that sentences are true.

Russell was the first antirealist to address this task seriously and he gave some very convincing answers. In his dispute with Meinong, Russell initially turned out to be the only winner. Antirealism about fictional and other nonexistent entities became a paradigm within the analytic philosophy and Meinong's thesis was considered definitely defeated by the objections moved by Russell[13]. At this regards, still in the seventies a philosopher like Ryle commented: “let us frankly concede from the start that *Gegenstandstheorie* itself is dead, buried and not going to be resurrected” (1973: 255). Not only Russell emphasized the apparent logical contradictions deriving from the application of the Independence Principle, stated by Meinong[14], and argued in favor of ontological parsimony, in accordance with the methodological principle known as the “Ockham's Razor”[15], but also, with his theory of

[12] See Everett (2003).

[13] At this regard, see Parson (1980, ch. 1). However, as Orilia (2002: 5) observes, it would be more correct to say that, rather than a Russellian paradigm, there was a paradigm intertwining, in various ways, both Russell's and Frege's positions. In fact, Frege's view has never been completely abandoned.

[14] See ch. 1.6.

[15] “Do not multiply entities beyond necessity” (*Entia non sunt multiplicanda praeter necessitatem*).

definite descriptions (1905), he provided an efficient method for the analysis of non-referring expressions that was adopted by antirealists for a long time[16].

Russell's eliminative strategy was to replace every problematic term with an adequate definite description, which is in turn logically analyzed in order to obtain a paraphrased version of the original sentence, where any reference to nonexistent entities has disappeared. Consider the sentence[17]:

3) Apollo is young

According to Russell, "Apollo" is not a genuine proper name[18], but it stands for a description expressing the information that we associate to a certain mythological character, namely the Greek god of sun, arts, poetry, and so on. Now, it is far from being obvious that such a move can be made, as we will discuss in chapter 2, but we can admit it for the sake of argument. Thus, we translate the name "Apollo" with an adequate definite description, let us say, "the sun-god". Then, following Russell's analysis, we paraphrase (3) into another sentence in which the description is eliminated:

(3R) There is at least one sun-god and at most one sun-god and every sun-god is young.

This strategy has the merit to give an account of sentences with empty names which is compatible with the standard logic, according to which sentences are either true or false. According to Russell, a sentence like (3), paraphrased in (3R), far from being meaningless, turns out to be false, since the conditions of existence expressed by the first part of the sentence – "there is at least one sun-god and at most one sun-god" – are not satisfied. As a consequence, the (external) negation of (3R):

[16] For a historical reconstruction of the Russell/Meinong debate, see Orilia (2002: ch. 8) and Reicher (2005).

[17] The example is taken from Voltolini (2010: 90).

[18] In later works (1910), Russell holds that every name of the ordinary language is synonymous for some definite descriptions. This is the descriptivist view that we will address in the next chapter.

(N3R) *It is not the case that* there is at least one sun-god and at most one sun-god and every sun-god is young.

is true. It follows that such a perspective also provides a straightforward way to treat negative existentals, like

4) Apollo does not exist

Instead of giving to (4) a Meinongian-style reading that presupposes the existence of what we deny, namely that there is a such-and-such entity but it does not exist, we paraphrase the sentences as:

4R) *It is not the case* that there is at least one sun-god and at most one sun-god.

Sentence (4R) is true because the definite description is not satisfied by any entity. Unfortunately, what Russell thought to be the greatest strength of his theory, appears to modern philosophers as its worst limit. The Russellian strategy is, so to say, too much strong, since it turns out that a considerable amount of sentences with empty names are false, whereas we would keep them as true[19]. Just consider sentence (3) again: leaving aside the Russellian paraphrase, our intuition it that (3) is true, not false, contrary to sentence:

5) Apollo is a detective

However, according to the analysis we have presented above, both (3) and (5) are false as well.

Some philosophers have tried to overcome this problem and save Russell's eliminativist method with paraphrases that are more compatible with common sense. Albeit

[19] For some philosophers, the Russellian strategy is even too much weak, since the paraphrase is essentially neutral with respect to ontological commitments, since it does not entail, but also it does not exclude, reference to nonexistent entities. See Kaplan (2005: 975–6) and Voltolini (2006: 139).

this strategy counts Lewis (1978) among its inspirers, the so-called *intensionalist* approach has mostly been defended by antirealists, like Rorty (1982), Lamarque & Olsen (1994), Orenstein (2003). They claim that the apparent truth of (3) is due to the fact that it corresponds to what the Greek myth tells us about Apollo, whereas (5) is false, from any point of view, because Apollo is nowhere regarded as a detective. We can, thus, accommodate our intuitions if we add, along with the Russellianparaphrases, a locution like "according to the Greek myth", working as an intentional operator. Thus, (3) is understood as elliptical for:

3I) According to the Greek myth, Apollo is young.

Which is, in turn, converted in

(3IR) According to the Greek myth, there is at least one sun-god and at most one sun-god and every sun-god is young.

The intensional operator has the function to shift the circumstance of evaluation of the sentence, in a similar way as modal operators do. While modal operators shift the circumstance of evaluation from the actual world to possible worlds, the intentional operator shifts the circumstance of evaluation from the actual world to the relevant fictional world:

> [The paraphrase] may be interpreted in terms of a possible-world semantics, according to which a sentence of the form "Iop," where "IO" is an intensional operator, is true iff its embedded sentence "p" is true in a (possible) unactual "world;" in our case the "world" of the fiction S. The word "possible" is put here in parenthesis because it may be reasonably asked whether the "world" of the fiction is a *possible* world. [...] Yet a defender of the present interpretation—let me describe it as the *genuine intensionalist* approach to sentences involving fiction—may limit him- or herself to taking the "world" of a fiction to be an unactual *circumstance* of evaluation for the embedded sentence. Accordingly, a sentence of the form "in the story S, p" is true iff the embedded sentence "p" is true in the "world" of the fiction S in question. (Voltolini 2006: 143-144)

This kind of approach also has the advantage to distinguish between sentences concerning different fictions, like:

6) Roland is wise

7) Roland is insane

They cannot be both true *tout court*, but they are if we consider (6) as related to the *Chanson de Roland* and (7) as related to the *Orlando Furioso*. By this way, we can also draw a comparison between how a certain character, Roland, is featured in a particular work of fiction and how the same character is featured in another one:

8) In *Orlando Furioso* Roland goes insane, while in the *Chanson de Roland* he is very wise.

Even more important, this conceptual apparatus allows us to distinguish between two kinds of sentences involving fictional names. From one side, we have *fictional sentences*[20], like (2), (3), (5), (6), (7) and (8), namely sentences that could easily occur in the body of a narrative text, like a fictional story or a myth. Such sentences say, from an inner perspective, what is true or false in the world the fiction is talking about and have merely fictional truth-conditions. On the other side, we have *internal metafictional sentences*, that talk about a fiction from a perspective that is external to the fiction itself. These sentences, such as (3I) or (3IR), are generated starting from fictional sentences by adding an intentional operator like "in the story S", or "according to the story *S*", and they have genuine truth-conditions: in fact, as said, a sentence of the form "in the story *S*, p" is true iff the embedded sentence "*p*" is true in the world of the fiction *S* in question.

The main problem[21], for the intensionalist account, is that not all sentences containing empty names can be paraphrased by means of an intentional operator. There are, in fact,

[20] See Voltolini (2010: 100-101) and Kroon & Voltolini (2016)

[21] There are many over critics that can be moved to the intensionalist approach, see Voltolini (2006: ch. 5; 2010: 97-114).

sentences about fictional characters that do not mention, even indirectly or implicitly, any particular pieces of fiction. We call them *external metafictional sentences*[22]. An example is:

9) Apollo is a mythological character

This sentence does not mention a property that Apollo has in some fictional world. At the beginning of the *Iliad*, the plague afflicting the Achaeans is not caused by a mythological – thus, nonexistent – entity, but by a powerful and angry god. Nonetheless, (9) is true, precisely because it tells us what Apollo is, over and above any specific fiction. Now, we would get the wrong result if we paraphrased (9) in

9I) According to the Greek myth, Apollo is a mythological character

The sentence is clearly false because, according to the Greek myth, Apollo does exist and he is, so to speak, a flesh-and-blood god.[23]

One could be more precise and distinguish, rather than between kinds of sentences, between different uses of a sentence containing fictional names. According to Evans (1982), a *conniving* use is when the sentence is uttered with the purpose to tell, within the perspective of a certain fiction, what happens in the world of that fiction. By means of such use, the sentence has merely fictional truth-conditions and does not entail any serious commitment to fictional entities. On the other hand, we have a *non-conniving* use when the sentence is uttered from a perspective that is external to the fiction, with the intention to express genuine truths that go beyond the context of the fictional narrative. Speaking of uses allows us to be more accurate in the analysis of sentences like:

[22] See Voltolini (2006: 172; 2010: 107). According to philosophers like van Inwagen (1977), Schiffer (1996), Thomasson (1999, 2003b) and Kripke (2013), external metafictional sentences force us to accept ontological commitment to literary and mythological entities. An alternative, but equivalent, terminology for fictional, internal metafictional and external metafictional sentences is the one proposed by Bonomi (2008): *fictive*, *parafictive*, *metafictive sentences*.

[23] A possible answer to this problem is to invoke a kind of fictionalism about fictional characters, by replacing the intentional operator "according to the story S" with one such as "according to the realist's hypothesis" (See Brock 2002 and Phillips 2000). But this seems rather an *ad hoc* move and has never really been adopted by antirealists. For a discussion on the topic, see Voltolini (2010: 109-112) and Voltolino & Kroon (2016).

(10) The Father is a fictional character

In fact, as Voltolini claims:

> Because there are metafictional stories [for example, Pirandello's *Six Characters in Search of an Author*] where some characters are directly described as fictional instead of being concrete individuals of some kind, we can have either a conniving or a non conniving use of an external metafictional sentence. (Voltolini 2010: 108, *my translation*)

Sentence (10) can be uttered from the inside of Pirandello's drama, thus with a conniving use, or can be uttered from the outside of the drama to feature the character with a property that only accidentally it has also in Pirandello's story itself. Similar cases occur with phenomena of "fiction within the fiction", or "theater within the theater", whose a famous example is provided by Shakespeare's *Hamlet*.

1.3.2 Pretense Theory

The distinction between conniving and non-conniving use enables us to introduce a new field of debate that has developed extensively in recent years. What does it mean, in fact, to utter a sentence from a conniving perspective? According to the so-called *pretense theory*, we utter a sentence with a conniving use when we are engaged into a game of make-believe, in which we put the reality aside and we imagine the situation depicted by the game, by accepting whatever is told *as* true[24]. Following Walton (1990), literary works are the outcome of games of make-believe, when they take the form of storytelling processes[25]. It follows that sentences that make up the body of a narrative, namely the fictional sentences, do not carry serious

[24] On the notion of make-believe, see Walton (1990) and Currie (1990).

ontological commitments, but only pretended ones. In fact, they can be evaluated only within the context of the relevant pretense. Of course, within that context, there are (typically concrete) individuals, like Sherlock Holmes or Siegfried. Yet, outside that context, there is no such individual at all. As Voltolini observes:

> Over and above the fiction that such a, typically concrete, individual exists, there is no further entity that the make-believe process is concerned with, not even things such as fictional entities. Put differently, the fiction that such a concrete individual exists not only does not assign a new, typically concrete, individual to the overall realm of the objects of discourse, but it also does not enrich that realm with an entity, say a fictional one, that allegedly does not exist in the same sense as a concrete individual does, that is, in the spatiotemporal way. There is a gap between making believe that there is a (typically concrete) entity and committing oneself to a fictional entity. The fact that, within the circumstance postulated by means of the relevant game of make-believe, there is indeed a (typically concrete) entity says nothing about whether there is a fictional entity existing outside that circumstance. (Voltolini 2006: 77)[26]

Consider, to resume a previous example, sentence:

2) Sherlock Holmes is a detective

Within the fictional world depicted in Holmes stories, there is an individual called "Sherlock Homes" and, since this individual is a flesh-and-blood man that helps people by brilliantly solving criminal cases, (2) is true about that fictional world. But the truth-value is merely

[25] As Evans (1982: 358) points out, make-believe games are fundamentally of two kinds: *existentially conservative games*, in which one pretends of certain really existing individuals that they possess such and such properties, and *existentially creative games*, in which one pretends that there are certain individuals possessing such and such properties. A similar idea is also present in Walton's distinction between a prop-oriented and a content-oriented make-believe games (See Walton 1990, 1993, 2000). The following analysis must be intended as concerning existentially creative games. We will discuss on existentially conservative games in ch. 3.

[26] However, Voltolini continues, even if we cannot take the existence of make-believe practices as a sufficient condition for the existence of fictional entities, this does not prevent philosophers who endorse a realist perspective to take them as a necessary condition for the existence of fictional entities. In other words, the notion of make-believe game is ontologically neutral: it neither excludes or entails the existence of fictional objects.

fictional. In fact, if we utter (2) with a non-conniving use, i.e., according to the perspective of our actual world, the sentence has no truth-conditions, for there is no individual the name "Sherlock Holmes" refers to.

At this point, the antirealist may try to save the truth-value of (2), even in its non-conniving use, by combining the pretense theory with the intentionalist approach[27]. So she can claim that (2) is elliptical for an internal metafictional sentence, containing the intentional operator "according to the story *S*". This attempt has the advantage to avoid the limit of the Russellian strategy, which is compromised with a descriptivist account of names[28]. In fact, the antirealist can hold that the locution "according to the story *S*" works not simply as a circumstance-shifting operator (like in the standard intentionalist theory), but as a *context-shifting* operator:

> That is, [the antirealist] may insist that it is an operator that shifts not only the circumstances of evaluation of the sentence it embeds, but also the context relevant for the interpretation of such sentence—typically, the context of its utterance. More precisely, if we take a fictional sentence "*p*" on its non-conniving use as elliptical for "According to the story *S*, *p*", then "*p*" so understood is true in the actual world if and only if "*p*", taken as uttered in the context of the story *S* (that is, a fictional context), is true in the world of that context. The antirealist merits of this account are clear. It allows a proper name like "Apollo" to be both genuinely empty, carrying no commitment to any fictional entity, but also genuinely non-descriptive. The embedded sentence containing the name is understood as being uttered in a fictional context, and in that context the name directly refers to an individual, the individual existing in the world of the relevant pretense. Since this reference occurs only in that fictional context, not in a real context, the name really does remain empty. (Voltolini & Kroon 2016)

However, the idea of a context-shifting operator faces some problems, as Kaplan, one of the most fierce critics of this view, has pointed out. In particular, context-shifting operators are in trouble when it comes to analyze sentences containing indexicals, for "no operator can control […] the indexicals within its scope, because they will simply leap out of its scope to

[27] For this proposal, see Walton (1990), Adams et al. (1997), and Recanati (2000).

[28] For a discussion on the limits of the descriptivist account, see ch. 2.5.

the front of the operator" (1977: 510). Consider the famous opening of Proust's *À la recherche du temps perdu*[29]:

11) For a long time, I used to go to bed early.

The first person pronoun "I", when (11) is uttered with a conniving use, i.e when it is evaluated within the fictional world of *À la recherche du temps perdu*, refers to the fictional storyteller of the romance, a person that does not exist in the actual world, no matters how many similarities he shares with the real Proust[30]. So far, so good. But if we want to capture the meaning of (11), when uttered with a non-conniving use, as being elliptical for a complex sentence introduced by the context-shifting operator, we obtain the wrong result. In fact, sentence

11 I) According to *À la recherche du temps perdu*, for a long time I used to go to bed early

should be true, if we follow the intensionalist analysis, since the embedded sentence "for a long time I used to go to bed early" is true in the world of *À la recherche du temps perdu*, as we have previously seen. But this is clearly a wrong reading of (11 I), for the intensional operator does not seem to shift the context of utterance for the embedded indexical "I". Contrary to what happens when (11) is uttered with a conniving use, the first person pronoun does not refer to the fictional storyteller of the "*Recherche*", but to the subject that actually utters the sentence, namely a subject existing in the real world, not in Proust's fictional one[31]. From this, it follows that (11 I) is false, unlike what is predicted by our previous analysis.

[29] The example is taken from Voltolini (2006: 166-7).

[30] For this point, see Bonomi (1994: 14)

[31] This point, as Bonomi observes, is controversial. In fact, there are other examples in which indexicals appear to shift their reference. Consider the sentence "In *War and Peace*, Napoleon Bonaparte is an arrogant person": in this case, "the use of the present tense [...] does not entail that the eventuality at issue (i.e., Napoleon's being an arrogant person) is a present eventuality" (Bonomi 2008: 25). For similar considerations, see Predelli (2008).

Some philosophers of the pretense theory have given up the intentionalist approach in favor of an alternative one, which reverses the order of explanation:

> Rather than taking the fictional sentence on its non-conniving use as elliptical for an internal metafictional sentence, we might take the internal metafictional sentence to be really true just when the fictional sentence (on its non-conniving use) is really true. [...] It is in fact relatively easy to discern the sense in which the fictional sentence on its non-conniving use is really true. It is really true just in case there is a pretense of a certain kind relative to which the sentence on its conniving use is fictionally true.
>
> (Voltolini & Kroon 2016)[32]

In short, as Crimmins claims, "the utterance [of fictional sentence] is really true just in case it is true within the pretense" (1998: 4). This account can be applied to both fictional sentences and external metafictional sentences. The latter, in fact, do not enjoy any special status: the only difference relies on the kind of game of make-believe that justifies the truth-value of the sentences. At this regard, Walton[33] distinguishes between *authorized* games, in which external constraints (*props*, in Walton's terminology[34]) like written or oral texts determines what is true or false in the world of the pretense, and *unofficial* games, which lack such external constraints. Unofficial games must be considered, as Recanati (2000: 226) points out, as generating a sort of Meinongian pretense, in which real individuals cohabit with fictional entities, the latter having, in addition to properties ascribed to them by literary works, also metafictional properties like "being a fictional character"[35]. External metafictional sentences are precisely based on games of this kind:

> [To say that an] external metafictional sentence (EMS) is really true is to say that what is really true is a sentence of the form "One who engages in a pretense of a certain kind (that is, one fictionally asserts EMS) within an unofficial game of make-believe of a certain sort

[32] The proposal was originally presented by Walton (1990) and then refined by Crimmins (1998).

[33] See Walton (1990: 51, 406, 409).

[34] See Walton (1990: 39–40; 1993, 2000).

[35] Similarly Crimmins, in order to analyze sentences that combine names of real and fictional individuals, like in "Ann is as clever as Holmes and more modest than Watson", introduces the notion of *shallow pretense*: "a shared, conspiratorial make-believe that we can refer with the names "Holmes" and "Watson" to people who are as described in Conan Doyle's stories." (1998: 3).

> (that is, one in which certain principles of generation are operative) is fictionally speaking truly". Put more simply, this amounts to saying that an external metafictional sentence is really true iff it is fictionally true when mock-asserted, namely iff in the "world" postulated by the relevant unofficial make-believe game things stand as the sentence fictionally says they do. (Voltolini 2006: 178)

So, for instance, the previous example:

9) Apollo is a mythical character

is true, on its non-conniving use, because there is a game in which (9) is fictional true: namely, there is a game according to which Apollo is a mythical character. Similarly, sentence

12) Donald Duck is loved by children all over the world

is fictionally true, hence also true on its non-conniving use, because there is a game in which, in addition to real children, there is a fictional character named "Donald Duck" so that it is loved by children all over the world. This is an unauthorized game, for no Disney comics features Donald Duck in such a way[36].

Even this account of fictional and metafictional sentences is not exempt from critics. The first objection is that the sentence's fictional truth within a pretense is not a sufficient condition for that sentence to be really true when used in a non-conniving way (it is, at most, a necessary condition): "no truth for a sentence in a "world" of fiction, so that the sentence is taken as having a certain meaning in that "world," can make that sentence true in the real world" (Voltolini 2006: 178)[37]. One can prevent the objection by adding some further conditions and by providing some appropriate antirealist paraphrases for metafictional sentences[38], but it is doubtful whether such move is applicable, in general, to all metafictional sentences. Anyway, the most controversial aspect of the pretense theory lies elsewhere, and

[36] See Voltolini (2010: 121).

[37] For this critique, see Voltolini (2006: 178-181; 2010: 122-3) .

[38] On this line, see Crimmins (1998).

precisely in the very notion of unofficial game, or shallow pretense (in Crimmins' terminology). This notion seems to rule out a crucial component of fictions, namely the element of meta-representation: when we make-believe that *p*, it is not sufficient to entertain a representation of *p*, but we also need a meta-representation of the action of making-believe *p*[39]. In other words, we must be conscious not only of the situation represented within the fiction, but also that such a situation is fictional. For instance, if a spectator, during the performance of a Shakespearean drama, jumps on the stage to prevent Othello from killing Desdemona, we would say that he does not know where he is or that he does not understand how theatrical performances work. The same we can say, *mutatis mutandis*, for the audience of the Lumière brothers' movie, *L'Arrivée d'un train en gare de La Ciotat*, if, as the well-known legend tells us, people were so terrified by the image of the train coming at them that they ran away from the cinema. Now, it is clear why the ideas of unauthorized game and shallow pretense are problematic: they do not respect the constraint of awareness required for the right understanding and appreciation of fictions. In fact, when we utter metafictional sentences like (9) or (12), we have not the intuition to be engaged in some sort of game of make-believe. But then, how there can be pretense, no matters how unauthorized or shallow it may be, if people are generally unaware of it?

To sum up, the antirealist approach, in none of the various forms it has taken, succeeds in providing a fully satisfactory answer to all the problems raised by the nonexistent. It is now the time to investigate the opposite view, i.e., the realist conception of fictional entities.

[39] The idea comes from Leslie (1987) and has been widely debated both in philosophy and in cognitive sciences. For a discussion, see Voltolini (2010: 46-50) and Meini & Voltolini (2010).

1.4 Realistic perspectives on fictional objects

1.4.1. Possibilist Theory

According to possibilist theories, fictional objects do not exist in the sense that they are not part of our world – the actual world – but nonetheless they exist, as *possibilia* that are located in some possible world[40]. The idea at the base is that, for whatever event narrated as a fictional story in our world, there is at least one possible world in which that event subsists as a state of affairs. In such world, the story is no more fictional, but told as known fact (Lewis 1978). By this way, we account for the intuition that fictional characters have the properties that characterize them in a given work of fiction. So, for instance, Sherlock Holmes is a human being that exists only in some possible world and has all the properties that Conan Doyle's stories ascribe to him: he is a brilliant detective, he has a friend called "Watson", a bitter enemy called "Professor Moriarty", and so on. Moreover, since fictional characters are usually depicted, in the stories in which they appear, as concrete entities with spatiotemporal existence (like people, animals, places, etc.), they are concrete entities in the worlds in which their story is true.

Possibilists theories must deal with many objections. One, moved by Kripke[41], is the so-called problem of ontological indeterminacy. It wonders how we can choose the right Sherlock Holmes among the many (maybe infinite) possible concrete individuals that possess all the properties ascribed to Sherlock Holmes in Doyle's stories, but that differ in other crucial ways (height, weight, eye colors, number of children, etc.). Moreover, nothing prevents that a single possible world contains many distinct individuals, each of them

[40] An historical, alternative, possibilist account of fictional and, in general, nonexistent objects is the one provided by Russell in the *The Principles of Mathematics* (1903), that I will not take into account. For a discussion of this topic, see Orilia (2002: ch.7).

[41] Kripke (1980). See also Kaplan (1973: 505–6).

satisfying all the relevant Sherlock Holmes properties. To these objections, different answers have been given. David Lewis[42], for example, can use his counterpart theory to offer a way to individuate and select each Holmes-candidate. According to him, every object exists in just one particular world (in Lewis' terminology, objects are "world-bound"), but it may have counterparts in other possible worlds. It may, because possible worlds do not necessarily have the same number of objects (in technical terminology, this is a *variable domain* conception of possible words). Fictional entities are indeed individuals that lack of a counterpart in our actual world. Thus, according to Lewis, since each Holmes-candidate is bound to just one possible world:

> Each Holmes-candidate is a counterpart for you of every other Holmes-candidate. For even if they should differ substantially in terms of overall qualitative similarity, the various Holmes-candidates are all counterparts by acquaintance for you (or, as Lewis seems to have thought, for your community of fellow-readers) – they are all, in their respective worlds, the person called "Holmes" whom you or your community (or rather, your counterparts) learn about by reading the Holmes stories, told as known fact.
>
> (Voltolini & Kroon 2016)

In this way, Lewis accounts for the plurality of possible Holmes-candidates by refusing to identify Holmes with any of them. The fictional character is indeed not a single possible object, but a set of *possibilia*, each one located in a different world, that are linked together by means of an appropriate relation of similarity[43].

Another objection that has been moved is that fictions are not limited to what is metaphysically possible: some stories contain more or less serious contradictions or describe situations and characters that are incompatible with reality. We must distinguish, here, between impossibility in a weak sense (venial impossibility), i.e., whose inconsistencies that arise by mistakes of the storyteller, and impossibility in a strong sense (blatant impossibility), i.e., whose inconsistencies that are a central and ineradicable feature of the story. The first case is easier to address within a possibilist framework[44]. A famous example is provided by

[42] See Lewis (1978, 1983, 1986)

[43] For a discussion on this topic, see also Currie (1990: chap.2), Kroon (1994) and Sainsbury (2010: 82–3).

[44] See Voltolini (2006: 12-13)

Watson's wound that "moves" from the leg (when you read *A Study in Scarlet*) to the shoulder (when you read *The Sign of the Four*). This is not a genuine contradiction, but rather we have two incompatible properties attributed to the same wound. According to Lewis, we can split Watson-candidates into two separate sets of *possibilia*: in the first set, there are those counterparts injured at the leg, in the second set, those counterparts injured at the shoulder.[45] On the contrary, cases of blatant impossibility, for instance a story that beings with the description of a wooden steel cannon[46], are instead much more difficult to treat, especially for who, like Lewis, does not make room for metaphysically impossible worlds[47].

Even more decisive, for Lewis's theory, seems to be another objection, again from Kripke:

> The mere discovery that there was indeed a detective with exploits like those of Sherlock Holmes would not show that Conan Doyle was writing *about* that man; it is theoretically possible, though in practice fantastically unlike, that Doyle was writing pure fiction with only a coincidental resemblance to the actual man. (See the characteristic disclaimer: "The characters in this work are fictional, and any resemblance to anyone, living or dead, is purely coincidental"). Similarly, I hold the metaphysical view that, granted that there is no Sherlock Holmes, one cannot say of any possible person that he *would have been* Sherlock Holmes, had he existed. Several distinct possible people, and even actual ones such as Darwin or Jack the Ripper, might have performed the exploits of Holmes, but there is none of whom we can say that he would have been Holmes had he performed these exploits. For if so, which one? (Kripke 1980: 157-8)

If we discovered that a man has lived who did all the actions ascribed to Sherlock Holmes, and even bore his name, available options would be only two: either we maintain that Sherlock Holmes is a fictional character, by saying that such similarities are merely coincidental, or we doubt that Doyle's stories truly are a figment of his mind. In the first case, we cannot identify the fictional character with the man of our example; in the second case, Sherlock Holmes is no more a fictional character and his stories are no more a work of fiction, but a trustworthy biography. As Kripke puts it, in no way we can say that similarity between

[45] See Lewis (1978: 45–6) and (1983: 277–8).

[46] The example comes from Twardowski (1894: 106) and it is also used in Voltolini (2006: 13).

[47] , However, some philosophers question this idea. See Yagisawa (1988).

an actual object and a fictional object overcomes the gap between fiction and reality. Then, moving from the actual world to possible worlds, things do not change: a possible entity, independently of how many properties it shares with a fictional object, cannot be that fictional object. As the gap between fiction and reality is unbridgeable, so it is the gap between fiction and possibility[48].

These critics lead many philosophers to put aside Lewis' account in favor of alternative versions of the possibilist theory. Some, and Kripke is among them, believe that many problems can be avoided if we adopt a semantics for modal logic based on a *fixed domain* conception of possible worlds, according to which there is only one single domain of quantification that contains all possible objects[49]. Since the domain does not change from world to world, contrary to Lewis' variable domain conception, the same individual can be a fictional nonexistent character in the actual world and an existent entity in other possible worlds[50]. This perspective accounts for the intuition that an author, for instance Conan Doyle, writes his novels because he has in mind one particular individual, namely Sherlock Holmes, rather than because he randomly chooses one Holmes-candidate from among all other possible ones. The individual "Sherlock Holmes" is such that it does not exist in the actual world but, in other possible worlds, it exists and does all the actions Conan Doyle tells us. Therefore, the fixed domain conception makes disappear the problem of ontological indeterminacy[51]. Moreover, at least according to Priest (2005), following this perspective we can also address cases of blatant impossibility, like the story of the wooden steel cannon, by admitting worlds that are incomplete and inconsistent.

However, difficulties for possibilist theories, whether they adopt a fixed or a variable domain, are not ended. As Crane observes, by holding that fictional characters are possible objects that have all the properties by which they are featured in a given piece of fiction, we obtain the counterintuitive result that fictions are *literally* true: "They are true of some possible world, not of the actual world; but they are true nonetheless" (Crane 2013: 76). But,

[48] Kaplan (1989: 609) had a similar intuition.

[49] A first defender of this thesis can be identified in Wittgenstein (1921).

[50] See Kripke (1980), Priest (2005), Berto (2011).

[51] However, analogous problems are raised. For example, since in the actual world there are many other possible individuals that do not exist but realize the Holmes stories in other possible worlds, one may wonder how is that Conan Doyle picked out just a single individual via the intentionality of his thought, as Priest claims, rather than any of the other actual Holmes-candidates. See Voltolini (2009) and Voltolini & Kroon (2016). For an answer to this objection, see Priest (2005: 93-94).

if works of fiction turn out to be trustworthy descriptions of some non-actual states of affairs, they can no longer be considered as products of a creative act of imagination, i.e., something that is made up rather than being just a report of facts:

> It is because authors are the creators of fictions that we value them as artists. Creating a fiction is a matter of making something up that did not really happen. The picture suggested by the concrete object realist, by contrast, is that authors happen to alight on a true description of something that *has* really happened. [...] What is obscure in this view, however, is why the invention of the story should be related to the real concrete possibilities which render it true. How can creating a story ("telling something as fictional") latch on to the wholly independent reality it describes?
>
> (Crane 2013: 77)

Finally, there is another serious problem. As Voltolini points out[52], since possibilist theorists hold that fictional objects do not actually have the properties that characterize them in the relevant story, it follows that fictional objects only have these properties in some of the worlds in which they exist[53]. So, for example, Sherlock Holmes actually is not a detective, but only possibly a detective. This position has to pay a cost of implausibility, in particular when we deal with sentences that compare fictional objects with actual concrete individuals. Consider the sentence:

13) Sherlock Holmes is cleverer than any actual detective.

The most natural way is to interpret (13) as saying that Holmes actually is cleverer than any other flesh-and-blood detective, since we are making the comparison in our own world, not in some other possible world. On the contrary, following Priest (2005: 123), we should read (13) as:

---

[52] See Voltolini (2010: 60-61) and Voltolini & Kroon (2016).

[53] I said "in *some* of the worlds in which they exist" for it is not true that a fictional entity, like Sherlock Holmes, is a brilliant male detective in all the worlds in which he exists: in some possible world Sherlock Holmes (or his counterpart) might have properties that differ from those ascribed to him in Doyle's stories. He could have undertaken a medical career, or be a woman, or be a totally inept detective, and so on.

13P) Sherlock Holmes, in possible worlds that realize the Holmes stories, is cleverer than any detective that exist in the actual world.

Such a reading seems to be odd and unjustified, because we normally do not read sentences involving a comparison between individuals in this way. Consider:

14) Stalin was crueler than any other actual dictator.

It would certainly sound bizarre to interpret (14) as saying:

14P) Stalin, in possible worlds in which Stalin fits the orthodox account of his activities, is crueler than any other actual dictator

Sentence (14) is intended to be a substantive claim about how Stalin actually was, not a claim about how he might have been. Moreover:

> a possibilist reading like [14P] makes it hard to make sense of the attitudes we hold towards fictional characters. Our belief in Holmes's great cleverness explains our admiration for Holmes, just as our belief in Anna Karenina's suffering explains our pity for Anna Karenina. It is hard to see how our belief in Holmes's possible intelligence could do this. Surely we do not get to be admirable merely by being thought to have admiration-evoking features in some other possible world. (Voltolini & Kroon 2016)

1.4.2 The Meinongian road

Meinong's theses became famous due to the debate that opposed him to Russell. While Russell defended the claim that there are no objects that do not exist, Meinong based his theory on the assumption that, besides objects that exist, there are objects that do not. As Orilia (2002: 4) observes, it must be noted that, even if Russell's claim may sound trivially true at first glance, from a purely logical point of view it is equivalent to the proposition, no longer so obvious, that everything that is an object exists, namely that everything exists[54]. Such a perspective would force us to exclude, along with many other things, fictional entities from our ontological catalogue. Thus, Meinong's approach deserves to be explored carefully.

Within the domain of what exists, Meinong (1904) distinguishes between two types of objects: on one side, concrete objects like tables, mountains, people; on the other side, abstract objects, like numbers and other mathematical entities. "Existence" (*Existenz*), in the sense of having a spatiotemporal location, is the kind of being that characterizes concrete objects. For abstract objects, Meinong introduces the notion of "Subsistence" (*Bestehen*). "*Sein*" is the word for the most general kind of being, which includes both subsistence and existence. Yet, over and above the domain of *Sein*, there are many other entities that have no kind of being. Such entities are "beyond being and non- being" (1904: 86).

In a similar way, Meinong distinguishes between two types of nonexistent objects: concrete and abstract objects. In dealing with them, I will concentrate only on concrete objects, which are generally taken as paradigmatic cases of Meinongian objects. On the one hand, there are those objects that fail to exist only contingently, in the sense that they could exist, even if actually they do not. They are possible objects, as it can be the golden mountain. On the other hand, there are those objects that fail to exist not only contingently, but necessarily[55]. By impossible objects we mean objects that have at least two incompatible properties, or a property that is the conjunction of two incompatible properties. A classic

[54] See also Quine (1953, ch.1).

[55] See Orilia (2002: 14).

example is the round square: it is a metaphysically impossible entity since nothing can exist that is round and square at the same time. Even more radical, from this point of view, are contradictory objects, namely objects such that, for some property *P*, exemplify both *P* and *non-P*, or a property that is the conjunction of *P* and *non-P*. A well-known example is "the square such that it is not the case that it is square"[56]. Lastly, there are incomplete objects which deserve special attention. By incomplete object we mean an object such that it is undetermined with respect to at least one property, i.e., for at least one property *P*, it exemplify neither *P* nor the negation of *P*[57]. Incomplete objects are necessarily nonexistent for they infringe the logical law of excluded middle:

> Pure things like the triangle as such are incomplete objects. The triangle as such has three angles (a constitutive property), it has three sides (a *consecutive* property, one that is somehow included in the object's *constitutive* properties), but it is neither scalene nor non-scalene, neither green nor non-green, etc. (Marek 2013)[58]

In general, we can consider fictional characters as examples of incomplete objects[59]. In fact, contrary to possibilist theories according to which each fictional character is a completely determined individual in the possible worlds in which it exists, according to (some) interpretation of Meinong's theory[60], fictional characters only have the properties that

[56] According to Meinong, contradictory objects infringe the law of noncontradiction with respect to internal negation only and not with respect to external negation. However, more recent philosophical works show that characterizing determinations can be constructed which go against any classical logical law, by also leading to contradictions involving external negation. For a discussion on this topic, see Marek (2013). Obviously, according to these definitions, all contradictory objects are also impossible objects, but not *vice versa*.

[57] "Incomplete objects violate the principle that for every property *P* and every object x, either x exemplifies P or x exemplifies the negation of *P*. However, incomplete objects do not violate the principle that for every property P and every object x, either x exemplifies *P* or x does not exemplify P" Reicher (2016). From this, it follows that a nonexistent object, like the golden mountain, in order to be a possible object must possess, besides the property of "being a mountain" and "being gold", an infinite number of other properties that completely determine it.

[58] Like contradictory objects, incomplete objects would infringe the law of excluded middle solely with respect to internal negation and not with respect to external negation. See Marek (2013).

[59] There may be exceptions, for example I can write a story in which the main character looks exactly like me except that he has blond hair instead of brown hair and he does exactly the same action that I do in my life.

[60] This is the most known interpretation on Meinong's view about fictional characters. However, his opinion on this point is not clear. Indeed, the theory that most plausibly is attributable to Meinong is a conception of fictional characters as higher-order entities, i.e., entities that are constructed out of simpler entities, in the

characterize them in the relevant stories. However numerous they may be, these properties are given in a finite number, whereas existing objects always have infinitely many properties:

> Note also that fictional entities so conceived are not completely determined with respect to their properties, unlike fictional entities conceived on the model of possibilism. Because Conan Doyle's stories are quiet on these matters, Holmes on the Meinongian model is not right-handed; nor is he left-handed; nor is he ambidextrous. He does, however, have the property of being one of these. (Voltolini & Kroon 2016)

Meinongian objects immediately give rise to problematic outcomes, such as our commitment to a seemingly paradoxical sentence such as "there are objects of which it is true that there are no such objects"[61]. To account for this perspective, Meinong relies on the so-called *Independence Principle*, which postulates a radical difference between "being" (*Sein*), i.e. for an object to have existence or subsistence, and "so-being" (*Sosein*), namely for an object to be in a certain way, i.e. to have some properties. The "so-being" of an object is the totality of the properties it has with the exception of its existence or non-existence. The Independence Principle precisely claims that an object may have any properties whatsoever, independently of whether the object exists or does not. This principle entails another one, which was explicitly formulated by Routley: "According to the *Characterization Postulate*, objects, whether they exist or not, actually have the properties which are used to characterize them, for example, where *f* is a characterizing feature, the item which *fs* indeed *fs*." (1980: 46)[62]. It follows that Meinongian objects have the properties by which they are characterized even though they do not exist. In other words, it is true that the golden mountain is both golden and a mountain, even if it contingently does not exist. Similarly, the round square is both round and square, even if it necessarily cannot exist. As regards fictional characters, according to standard interpretation, they are indeed Meinongian objects possessing the properties they are characterized in their stories, even if this makes them incomplete objects[63].

same sense in which, for instance, a melody is an entity constructed out of its constituent sounds. See Voltoloni (2010: 63-4) and Voltolini & Kroon (2016). For a discussion about this topic, see also Kroon (1992: 503–10), Raspa (2001), Barbero (2005).

[61] Meinong (1904: 83).

[62] The formulation of the Characterization Principle is already implicit in Meinong (1904: 82).

[63] See Voltolini (2006: 4-12) for a discussion on this point. In particular, as regards the problem whether Meinong adopts a phenomenological conception, according to which *ficta* are intentional objects generated

Such a perspective allows us to respond to some objections moved against possibilist theories, in particular the problem of ontological indeterminacy and the issue concerning the interpretation of sentences that compare fictional objects with actual concrete individuals.

Apart from the problems strictly linked to the fiction, philosophers have always been cautious in accepting Meinong's theory, due to the numerous paradoxes and puzzles that nonexistent objects arise. Not by chance, it was Russell's paradigm to prevail, at least initially. However, starting from the 60s and the early 70s, a group of philosophers recovered Meinong's thesis, by responding in various ways to Russell's critiques[64].

According to Voltolini (2010: 64), modern theorists inspired by Meinong can be divided into two groups: *orthodox* and *unorthodox* Meinongians. Philosophers of the first group[65] maintain that Meinongian objects are concrete correlates of sets of properties. In other words, for every set *S* of properties, for instance "being square" and "being round", there is a concrete object that has, at least, the properties contained in *S*, namely the round square. These properties constitute the essence of that particular object. However, a Meinongian object may have some other properties, for instance the property of "being often thought about by Meinong". As regards fictional characters, neo-Meinongian conception does not differ so much from what we have already said above. We consider the fictional story as defining a set of properties, to which it corresponds a concrete nonexistent object endowed with all those properties[66]:

> Fictional objects can similarly be regarded as concrete correlates of sets of properties on such a view. Consider Holmes, a cocaine-addicted detective who lives (lived) in London at 221B Baker Street, solves many baffling crimes, has a friend called "Watson", and so on (where the "and so on" includes all the properties P such that it is true in the Holmes stories that Holmes has P). For orthodox neo-Meinongians, Holmes is a concrete, albeit

by an act of imagination, or a Platonist conception, according to which *ficta* are mind-independently constituted objects.

[64] Beyond neo-Meinongian philosophers, there are also academics who have dealt extensively with the ontology of literary discourse that tend to propose theses quite close to Meinong's positions. For instance, see Bonomi (1994) and Eco (1997).

[65] See Parsons (1980), Routley (1980), Castañeda (1989a, 1989b), Jacquette (1989, 1996).

[66] Not all Orthodox neo-Meinongians accept this view. An original proposal within this field is represented by Castañeda's guise-theory (1989a).

> nonexistent, correlate of the set having those properties as members. (Voltolini & Kroon 2016)

Philosophers of the second group, whose the most representative member is Zalta (1983, 1992)[67], disagree with the orthodox view with respect to one fundamental issue: they equate Meinongian objects to abstract objects, namely to entities that lack spatiotemporal existence, but not existence *tout court*. From this it follows a Platonic reading of Meinongian objects, and *ficta* in particular, as *generic objects* or *roles*. Consider the sentence:

15) The US President faces an election every four years.[68]

Here, with the description "the US President" we are not referring to a particular Us President like Washington, Jefferson, Roosevelt or Lincoln, but we mean rather to talk about the role or office of the US President that those individual covered after having been elected. Similarly, when we talk about Sherlock Holmes we are not referring to a human being, albeit nonexistent, because human beings are not abstract objects. Instead, we are talking about something like the role for a theatrical drama or a movie that can be played by one actor or another[69].

As we said, the fundamental difference between the orthodox and the unorthodox views concerns the way they conceive Meinongian objects at the metaphysical level: as concrete nonexistent objects, form one side, as subsistent abstract objects, from the other side. Further differences can also be drawn, in particular with regard to their attitude towards the distinction between nuclear and extranuclear properties[70] and, in parallel, the distinction

[67] Among those philosophers who have a theory assimilable to Zalta's approach, we can also include Rapaport (1978).

[68] The example is drawn from Voltolini (2006: 65).

[69] Similar thesis are also supported by philosophers that do not directly drawn their inspiration from Meinong, such as Currie (1990) and Santambrogio (1992).

[70] A first distinction of this kind was sketched by Meinong himself (1916).

between internal and external modes of predication[71]. We do not have the possibility to address these issues in details[72]. It is sufficient to say that, despite the differences:

> Neo-Meinongianism in all its varieties tends to sketch a Platonistic picture of a fictional entity, either as something akin to a Platonic attribute, or as a correlate of something else that we tend to describe in Platonic terms – a set of properties. Neo-Meinongianism thus sees a fictional object as something that pre-dates the story-telling activities that intuitively bring fictional objects into being. (Voltolini & Kroon 2016)

The Platonist perspective derives from the adoption of the *Principle of the Freedom of Assumption* which, in its weakest formulation, postulates that for any set *S* of properties there is an object that shares those properties[73]. Regarding the field of fiction, this principle automatically guarantees that, for any story, there is a fictional character having the properties by means of which it is featured in that story. However, such a position may result in a weak point. Given this background, some objections arise that affect neo-Meinongian theories of both sides.

First of all, there is the "no-*ficta*" problem. An example is provided by the word "Moloch", which ancient interpreters of the Bible erroneously referred to a mythical pagan god. Conversely, modern philology has demonstrated that this word was used as a generic term either for kings or for human sacrifices[74]. Notwithstanding their error, there is undoubtedly a set of properties that former interpreters mistakenly understood the Bible to assign to a certain pagan god. The problem is that, given the principle stated above, we should have a Meinongian object corresponding to such a set of properties, even if in the Bible there is no reference to an entity of this kind[75].

---

[71] The distinction between modes of predication was originally proposed by Ernst Mally (1912).

[72] For an in-depth discussion on these topics, see Voltolini (2010 61-74) and Voltolini & Kroon (2016).

[73] See Meinong (1910: 246-7; 1916: 282; 1917: ch. 2). Over the years, Meinong admitted some limitations to the principle. For a discussion on this topic, see Marek (2013). To Meinong is also ascribed a stronger version of the principle, according to which for any set *S* of properties there is an object that has *exclusively* those properties. See Orilia (2002: 86-88).

[74] See Vololini (2006: 32), (2010: 73). The example is drawn from Kripke (2013).

[75] An attempt to prevent the "no-*ficta*" problem can be found in Castañeda (1975) and (1989b, ch.11), but with not very convincing results (see Voltolini 2010: 73). See also Orilia (2002: 148).

A similar, even more serious, difficulty is represented by the "many-*ficta*" problem. As Voltolini observes:

> It is not only the case that a set of properties does not by itself generate any *fictum*; it may also be that it matches more than one *fictum*. Take the famous example in which Jorge Luis Borges imagines that a man called Pierre Menard happens to write a text that is word for word identical with Miguel Cervantes' *Don Quixote*. Suppose one idealizes this case so that not only is Menard taken to be totally unconnected with Cervantes (neither knows anything about the other), but the two literary works mobilized by such texts are imagined to coincide not only in their explicit but also in their implicit truths. [...] In that case, one and the same set of properties matches different characters. If we take Borges' example as one of these cases, Cervantes' Don Quixote and Pierre Menard's Don Quixote are two distinct characters who, nonetheless, share all the properties attributed to them in the respective works. (Voltolini 2006: 32-33)[76]

Given that Cervantes' and Menard's *Don Quixote* share the same properties in the respective works of fiction and that Menard and Cervantes are unknown to each other, the example mentioned above describes a situation which leads us to an unwanted result: since by the *Principle of the Freedom of Assumption* there is a one-to-one correspondence between sets of properties and their correlates, we obtain a single Meinongian object whereas, according to our intuitions, *ficta* involved are, in fact, two.

Finally, we can move the same objection that has already been presented against possibilist theories. Since the Platonistic framework, common to both orthodox and unorthodox neo-Meingonian theories, urges us to conceive fictional entities as Meinongian objects that are given independently from any fiction, the work of a storyteller is no more the product of a creative act of imagination, but it is reduced to be a mere description of some actual states of affairs. As Thomasson points out:

> Whether fictional entities are taken to be unactualized possibilia, non-existent objects, or abstract kinds, it seems that in any of these cases the work of authors writing stories is completely irrelevant to whether or not there are these fictional entities: the relevant

[76] See also Voltolini (2003) and (2010: 74). It is an appropriate simplification of the original short novel *Pierre Menard, autor del Quijote* by Borges (in *Ficciones*), in which the protagonist not only writes a novel identical to Cervantes' story, but does so intentionally. The case of Pierre Menard was originally proposed by Lewis (1978: 39), and was also adopted by Currie (1990: 77-8), Thomasson (1999: 56) and Fine (1982: 107).

> possibilia, non-existent objects, and abstract kinds were "around" just as much before as after acts of authoring, and so we can't take seriously the idea that authors create fictional characters on any of these views. The best these views can do to account for the apparent truth of claims such as "Hamlet was created by Shakespeare" is to say that it is at least true that Shakespeare *described* or *selected* Hamlet from among all the available *possibilia*, non-existent objects, or abstract kinds and, by writing about that object, made it *fictional*. (Thomasson 2009: 15)

1.4.3. Artifactualist Theory

The intuition that fictions are created or made up is at the base of the last metaphysical theory on fictional objects that we are taking into account. This form of realism holds that in addition to creating fictions, authors also create fictional characters. So, for instance, Conan Doyle and Flaubert created, along with their novels, also the primary and secondary literary characters appearing in those novels, like Sherlock Holmes and Watson, or Emma Bovary and Léon Dupuis. Such authorial creations are not flesh-and-blood individuals, but abstract objects. However, unlike the unorthodox neo-Meinongian view, they are abstract objects of a particular kind. Following a philosophical tradition that goes back to Ingarden (1931) and Husserl (1948), we can divide abstract objects into two classes: *free idealities*, namely Platonic entities characterized by a timeless, changeless, necessary mode of being, and *bound idealities*, namely abstract objects that depend for their existence on the existence of other beings[77]. This dependence must be intended both metaphysically – in the sense that they cannot exist unless some other beings exist[78]– and temporally – in the sense that they are brought into existence at a certain time and they last as long as certain conditions are maintained. The class of bound idealities includes all those things that are the historical

[77] For a discussion on Husserl's and Ingarden's theories, see Smith (1980). For a general discussion on the notion of ontological dependence, see Fine (1994).

[78] Philosophers give a modal reading of such dependence: the dependent entity exists only in possible worlds in which those other beings exist. To be more precise, an entity *O* existentially depends on another entity *O′* iff, necessarily, if *O* exists, then *O′* exists. See Thomasson (1999: 36–7).

outcome of some cultural practice. Paradigmatic examples are institutional entities, such as constitutions, nations, governments, universities and, more generally, social and cultural objects as human languages, marriages, symphonies, and works of literature themselves. Within this category, we can also insert Strawsonian types, intended as historically established principles of construction of like particular (tokens):

> The general title of "types", often, though rather waveringly, confined to words and sentences, may well be extended. Works of art, such as musical and literary compositions, and even, in a certain sense, paintings and works of sculpture; makes of thing, e.g. makes of motor-car, such as the 1957 Cadillac, of which there are many particular instances but which is itself a non-particular; and more generally other things of which the instances are made or produced to a certain design, and which, or some of which, bear what one is strongly inclined to call a proper name, e.g. flags such as the Union Jack.
>
> (Strawson 1959: 231)[79]

Finally, according to some philosophers[80], also entities that are not the result of human activities, like biological species, can be considered example of bound idealities: in fact, a biological species is an abstract entity – we never encounter the species, but only particular individuals – that exists insofar as there are specimens of it: when, in the seventeenth century, the last specimens of dodo died, the species dodo (*Raphus cucullatus*) disappeared with it.

As far as fictional entities are concerned, the standard account is the following:

> A fictional character is an abstract cultural artifact created at a certain time by the acts of an author writing a work of fiction. Like other artifacts, it is a contingent member of the actual world—for example, if Arthur Conan Doyle's medical practice had been busier and he had never had the time to write, Sherlock Holmes might have never existed. But although created and contingent, fictional characters are also abstract entities – where "abstract" means just lacking a spatiotemporal location (although they may have certain temporal properties such as a time of creation), not being among the timeless, changeless, necessary, entities of the Platonist. They are thus relevantly similar to other abstract cultural creations such as laws and theories, works of literature and music, all of which fall between

[79] For a discussion on Strawson's perspective, see Wetzel (2008) and Thein (2009). An interesting attempt to apply the notion of "type" to fictional characters can be found in Terrone (2017a).

[80] See Künne (1982: 407)

> traditional bifurcations of categories into real (spatiotemporal) and ideal (Platonistic) entities. (Thomasson 2003a: 139-40)[81]

This is the so-called *artifactualist* or *creationist* theory. The doctrine was first defended by Ingarden (1931) and then by several scholars on different occasions[82]. In recent time, a more complete development of the artifactualist account – the one that we will take into consideration here – has been developed by Schiffer (1996, 2003), Salmon (1998) and, among all, Thomasson (1999, 2003a, 2003b, 2009)[83].

According to Thomasson, dependence of fictional entities must be understood in two ways. First of all, they *historically rigidly* depend on the authors who create them (necessarily, if *O* comes into being at the time *t*, then the author(s) who creates *O* exists at some time *t'* before *t*)[84]. As Voltolini observes:

> A *fictum* depends rigidly on the particular mental act of the author of the fiction that talks of it: if that act did not exist, the *fictum* would not exist either. Furthermore, this dependence is not only rigid but also historical, where by "historical dependence" Thomasson means that in order for the dependent entity to come into existence, another entity must already exist. Hence, a *fictum* also requires that the mental act on which it rigidly depends already exists in order for it to come into existence: had that act not occurred, the *fictum* would not have come into existence. Historical dependence is what entitles us to speak of the *fictum*'s original thinker, namely the author of the fiction talking of it, as its creator. (Voltolini 2006: 40-41)

Historical rigid dependence preserves, contrary to both possibilist and (neo-)Meinongian theories, the idea that fictional characters are created and contingent members of the actual world, albeit they have no spatial location. In fact, not only a fictional object begins to exist at

---

[81] Note that the idea that fictional entities are created by their authors is not so obvious as it may appear. For some objections on this point, see Yagisawa (2001) and Brock (2010).

[82] Van Inwagen (1977), Searle (1979), Kripke (2013).

[83] For an in-deep presentation of the Artifactualist theory, see also Vololini (2006, ch. 2) and Predelli (1997, 2002).

[84] Thomasson (1999: 31).

a certain time, but also it cannot exist unless a particular entity exists, namely the particular activity of an author writing a work of fiction.

By means of this conceptual apparatus, the artifactualist theory addresses the intuition that fictional characters are not entities that are given before and independently from any activity of fiction-making, but they are made up, or created, by the authors of the stories in which they appear: if, for instance, Arthur Conan Doyle had never written his novels, Sherlock Holmes would not exist. Moreover, by holding that fictional characters (like works of literature themselves) are created by the activities of authors and identified primarily by their historical origin, artifactualists may avoid many metaphysical difficulties affecting the theoretical perspectives we have illustrated before. In particular, the "no-*ficta*" and the "many-*ficta*" problems. As for the "no-*ficta*" problem, exemplified by "Moloch" case, it is clear why there is no such a *fictum*: no "Moloch" story is really contained in the Bible. Thus, there is not the act of an author which generates such a literary character by gathering certain properties and embedding them within the body of a narration. Of course, this does not amount to say that it does not exist any fictional character called "Moloch". Surely Moloch appears in some other narratives, but not in the Bible. And since Moloch has not been brought into existence by biblical authors, it must have some other creators, such as the interpreters of the Bible that erroneously supposed it to exist[85]. The solution of the "many-*ficta*" problem is similar. Even if we have two stories depicting a fictional character with exactly the same properties, we can distinguish between Cervantes' and Menard's Don Quixote because, ultimately, there are two independent acts of authorial generation[86].

According to Thomasson, dependence of fictional entities is not limited only to rigid historical dependence on a given mental act of its creator. Over and above this, fictional entities also *constantly generically* depend on literary works or other media that feature them (necessarily, if *O* goes on existing, then some literary work *W* or other featuring *O* exists during *O*'s time of existence)[87]. On this point, Voltolini observes:

---

[85] In this case, Moloch is no more an object of fiction, but an object of error. See ch. 1.1.

[86] See Thomasson (1999: 56). See also Voltolini (2006: 41; 2010: 78). Nonetheless, as Voltolini points out, it is not clear how artifactualist theory can deal with a related issue, namely the problem of *indiscriminable* fictional objects: intuitively, there were thousands of fictional dwarves who took part in Tolkien's War between the Dwarves and Orcs, without Tolkien's engaging in thousands of acts of dwarf-creation.

[87] Thomasson (1999: 36).

> In order for a *fictum* to exist, therefore, there must be a narration that speaks of that *fictum*. Moreover, this generic dependence is also constant. By "constant dependence" Thomasson means a relation such that the dependent entity requires that the entity on which it depends, exists at every moment at which the dependent entity exists. As a result, by being constantly and generically dependent on literary works, a *fictum* requires that, at every moment it exists, there is a literary work that mentions it. (Voltolini 2006: 42)

Constant generic dependence justifies the fact that a fictional character, besides being generated at a certain moment, persists through time and can also disappear. In order for Sherlock Holmes to continue to exist, we need the presence of the literary works featuring it, and we have literary works as long as there are material copies (tokens) of them[88]. If every copy of Holmes stories were destroyed, Sherlock Holmes, like the dodos, would disappear as well. Moreover, such kind of dependence accounts for the intuition that we may describe fictional characters as having a certain age: Sherlock Holmes, being created in 1887, is now over 130 years old. But even more interestingly, we can account for the intuition that literary characters may have more or less success, in the sense that they may appear in many works of literature such as sequels, remakes or parodies that describe them in different ways and sometimes bring them to even surprising developments. In fact, fictional characters can not only "migrate" from one work to another, but even transit across different media (from a book to a comic or a movie, from a text to a picture to a sculpture):

> [Artifactualist theory] also preserves the idea that one and the same character may appear in many related texts (sequels, parodies, and so on). Since character identity is based on historical continuity rather than sameness of properties, a single character like Sherlock Holmes may reappear in a number of different stories, for a later author may refer back to an extant character and ascribe it new properties, even if the properties ascribed are rather different than those ascribed to the character in the original work.
>
> (Thomasson 2003a: 140)[89]

This intuition is hard to explain if we adopt other approaches, especially neo-Meinongian theories, that typically individuate characters in terms of their associated properties so that a

[88] To be more precise, for Thomasson (1999: 11–2, 36) even a mental remembering of a work may suffice in order for that work to be kept in existence.

[89] See also Thomasson (1999: 67-69).

character cannot be literally said to appear in more than one story where so much as a single property is changed[90]. Whenever we move from one work of fiction to another and some features of a certain literary character are changed, a *fictum* replaces another *fictum*, no continuity is possible[91].

Even if artifactualist theory allows us to address many different intuitions, this perspective is not immune to critics. The main objections concern the nature of the creative process and the relation between the creative process and the identity of fictional objects. First of all, since Thomasson claims that the origin of a *fictum* is due to an author's act of imagination, it seems that what comes into existence is not, *ipso facto*, a fictional object, but rather an intentional object, i.e., the target of a certain authorial thought. If so, a fictional object is not such at the origin, but only at a later time. We must then specify what makes a merely intentional object to become a fictional object. The issue may be addressed by giving a somewhat different account of the generating process. For instance Schiffer (1996, 2003)[92] identifies the origin of a *fictum* as an abstract artifact not in the author's act of conceiving it in his mind, but at the end of the whole process of creating a fictional story. However, since this process is based on a practice of make-believe that has no ontological commitments, for the storyteller merely pretends that there is an individual with such-and-such properties, it is not clear how such a process should lead to the generation of independently existing objects like fictional entities[93]. But, maybe, the most serious difficulty regards the identity conditions of fictional objects. The problem arises because artifactualist theorists typically deny that fictional objects really have the properties by which they are characterized *within* the fiction. In fact, fictions generally depict literary characters as concrete individuals: in Doyle's stories, Sherlock Holmes is a flesh-and-blood detective who interacts with physical objects and other flesh-and-blood people. But Sherlock Holmes, as an abstract artifact, cannot be depicted in

[90] For a discussion on how neo-Meinongians can answer to this problem, see Reicher (1995) and also Eco (1979, 2000) and Orilia (2002, ch.14).

[91] As Voltolini observes: "To the extent that they do not share the same properties assigned to them in the course of the relevant narratives, the Roland of *Orlando Innamorato* is not [...] the Roland of *Orlando Furioso*. Of course, [...] all these fictional characters can be linked to a sort of *general* character: the Roland in general, the Faust in general, the Ulysses in general. But it is not clear how this general character should be featured." (Voltolini, 2010: 77-78, *my translation*).

[92] Also Thomasson (2003a; 2003b) modified her view by adopting Shiffer's proposal.

[93] For a debate on this point, see Voltolini (2009b).

such a way. And in fact, Sherlock Holmes does not *really* possess those properties. According to Thomasson:

> We can distinguish "fictional contexts" of discussion about what is true from within the context of a work of literature (or while involved in the pretense it demands of readers) from "real contexts" in which we discuss works of literature and their characters from the "real world" perspective. Statements about fictional characters within fictional contexts may be read roughly as statements about what is true *according to the story*, so "Holmes is a person" is to be read as implicitly claiming "According to the stories, Holmes is a person", and so understood of course is true. Real context claims such as "Holmes is a fictional character (and thus an abstract artifact)" can be read as straightforwardly true, removing the apparent conflict between them. (Thomasson 2003a: 141)[94]

It follows that only a few number of properties can be properly ascribed to fictional objects, namely those properties that are predicated of it from a perspective that is external to the fiction, like "being a fictional object", "being famous all over the world" or "being Doyle's creation"[95]. This approach has two problems: from one side, it fails to preserve a sense in which one can say that fictional objects *actually* have the properties characterizing them in the relevant stories, from the other side, this restricted set of properties makes it hard to see how to individuate a fictional entity. For instance, also Watson is a fictional character created by Conan Doyle and it is famous all over the world, but Watson is not identical with Sherlock Holmes[96].

As regards this problem, Thomasson gives a sufficient identity condition to individuate literary characters within a single work of fiction: *x* and *y* are the same fictional object *F* if *x* and *y* are ascribed exactly the same properties in the work[97]. But this condition cannot be applied for the identification of literary characters at the *intertextual* level, i.e., when the same character appears in different works of fiction. In this latter case, we can only

[94] See also Thomasson (1999: 111-114).

[95] In other words, the properties that are typically mobilized in external metafictional sentences.

[96] See Voltolini (2010: 81) and Voltolini & Kroon (2016). To be precise, Thomasson can answer that a fictional character also has properties like "being a detective according to the Holmes stories", that make it easier to identify it with respect to other fictional characters, like Watson.

[97] Thomasson (1999: 63). For a counterexample to the sufficient identity condition proposed by Thomasson, see Voltolini (2010: 82)

provide a necessary condition: *x* and *y* are the same fictional object *F* only if the author of the second work *W′* is competently acquainted with *x* of the previous work *W*, and intends to import x into W′ as y[98]. For instance, Sherlock Holmes in *A Study in Scarlet* is the same fictional character that appears in *The Sign of the Four* because the author, Conan Doyle, intends to import the character of the first book in what is, in practice, a sequel of *A Study in Scarlet*, and of course Conan Doyle is acquainted with the character he has invented. However, there are not also sufficient conditions because, as Thomasson admits, an author, despite his intentions, may not succeed in importing a fictional character into a new work of fiction if the features of the character in the new story are too far from the original ones. For instance, even if Gregory House, the main character of the TV series *Dr. House*, has an evident resemblance with the detective invented by Conan Doyle, we do not say that Gregory House and Sherlock Holmes are the same fictional character, no matter what the author's intentions are. At most, we can say that the latter character is inspired by the former one, and that is all. But, for the artifactualist theory, there are even more challenging examples in which authorial intention fails to import fictional characters, as cases of fusion and fission show out[99].

As a final remark, we can say that artifactualists and neo-Meinongians seem to be, in a certain way, complementary, since the former focus the attention on the existence conditions for fictional objects, whereas the latter on the identity conditions. Maybe, this can make room for the development of syncretistic accounts able to bring together the qualities of the two perspectives:

> neo-Meinongian and creationist theories seem to suffer from complementary defects. On the one hand, neo-Meinongians provide exact identity criteria for fictional objects, but these criteria are clearly insufficient in that they do not take into account the fact that such objects are products of the human mind. On the other hand, creationists do account for this fact, but they only provide relatively non-specific identity criteria for such entities. Those theories are normally taken to be incompatible, for they appeal to different metaphysical models—broadly speaking, a Platonic model as opposed to a constructivist one. This claim of incompatibility should not be taken as definitive, however; there may well be ways in

[98] Thomasson (1999: 67-8)

[99] We have a fusion when the author intends to import into *W′* two characters *x* and *y* from a previous work *W* as a single character *z*, whereas there is a fission when the author intends to import into W′ a character x from a previous work *W* as two characters y and z. These cases will be discussed in detail in the ch. 3.7.

> which the two theories, or perhaps the most promising elements of each theory, can somehow be combined. (Voltolini & Kroon 2016)[100]

In conclusion, we can say that, despite the numerous proposals that have been examined, neither the realist nor the antirealist approaches succeed in giving a fully satisfactory account of the nonexistent. Literary characters elicit various and sometimes contrasting intuitions that are very hard to address. It is maybe the time to tackle the topic starting from a different perspective. I will deal with some issues concerning the cognitive nature of fictional objects, i.e the way by which we conceive them in our mind. I think that it is a preliminary task that we should do before proceeding in the investigation about the ontological and metaphysical questions fictional objects pose to us. From my point of view, the theory of mental files provides us with the best framework to succeed. Next chapter is then devoted to present this theory and the background from which it originates.

[100] For some proposals of syncretistic theory, see Zalta (2000) and Voltolini (2006, ch.3).

CHAPTER TWO

THE THEORY OF MENTAL FILE

2.1 Millian view of singular terms and its challenges

According to the traditional view, we are able to represent states of affairs, i.e. facts, by means of ordered and meaningful sets of words, i.e. sentences. In the simplest cases, a state of affairs amounts to an object and a property. Sentences that we use in these circumstances are as simple as the facts one aims to represent: we just need one name for the object and one predicate for the property we assign to that object. For instance, if we utter a sentence like:

1) Socrates is a human being

we mean to predicate, of a certain individual called "Socrates", that he is a human being (with all that follows, including his short-term existence). Moreover, it seems that we are uttering something true about such an individual. Thus, a question immediately arises: how do we manage to talk about Socrates and thereby say something meaningful and even true about him? To circumscribe the field of investigation, we can focus on a more specific question: how do we refer to Socrates by means of the name "Socrates"?

Language provides us with some tools which have the function to pick out, or refer to, objects that are scattered in the external world. (Here, by "objects", I mean concrete spatiotemporal located entities, like people, animals, trees, mountains, tables, towns, and so on.) These tools are represented by the class of words and expressions that includes proper names, demonstratives, indexicals and definite descriptions. They are generally called *singular terms*, or *directly referential term*:

> A [direct] referential term is a term that serves simply to refer. It is devoid of descriptive content, in the sense at least that what it contributes to the proposition expressed by the sentence where it occurs is not a concept, but an object. Such a sentence is used to assert *of* the object referred to that it falls under the concept expressed by the predicate expression in the sentence. (Recanati 1993: 3)

In particular, proper names are generally regarded as paradigmatic examples of singular terms: "if there are such things as referential terms — that is, terms that somehow 'hook on to' things in the world — then proper names are surely among those terms" (Reimer & Michaelson 2017). For the purposes of this work, I maintain that proper names are, roughly speaking, co-extensive with the kind of expressions that ordinary speakers – the so-called man-of-the-street – standardly call "names": there are names for people, such as "Socrates", "Aristotle" or "Napoleon", names for places and cities, such as "Paris" or "Rome", names for natural or artificial objects, like "Mount Everest" or "Tour Eiffel". What do these expressions have in common? In virtue of how do they constitute a genuine class of linguistic expressions? John Stuart Mill is given credit[101] for the common view that the contribution of a name, at the semantic as well as at the cognitive level, is its referent, and nothing else. Thus, "Socrates" refers to a particular man, "Paris" refers to a particular city and "Mount Everest" refers to a particular mountain. But the Millian conception of names faces so many challenges that no one today holds this position, at least in its original version. A huge array of problems is indeed provided by fictional names. In fact, if one does not endorse a realistic approach about fictional entities, every sentence containing the name of a literary character, like "Sherlock Holmes", cannot express genuine truths or falsehoods about the world, for the name is empty and the sentence remains meaningless. But such a position is hard to hold, without adopting some sophisticated expedients, as we have already widely discussed in the first chapter[102]. Below, I shall briefly expose some other well-known problems, or puzzles, stemming from the Millian view, in order to better understand the request for alternative theories.

[101] This is a standard oversimplification. It is probably not exactly what Mill (1843) said, and in any case it is unlikely that Mill was the first one to hold this view. See Cumming (2016) and Sainsbury & Tye (2012: 21 – 26).

[102] See ch. 1.3.

2.2 Frege's Puzzle About Identity Statements

Let us mention an anecdote. Since the planet Venus is relatively close to the Sun, it is one of the most brilliant heavenly bodies observable from Earth. It can usually be seen only for few hours right after sunset, on the west side upon the horizon, or just before dawn at east. Such behavior of Venus may have generated misunderstandings among ancient astronomers. According to the well-known legend, Babylonians mistakenly supposed to see two separate bodies. Thus, they introduced the name "Hesperus" for Venus as seen in the sky at evening, the name "Phosphorus" for the same planet as seen in the morning. If we consider that some alternative expressions to designate Venus were "the Evening Star" and "the Morning Star", we can appreciate how deep was the confusion on that issue.

Of course, it is just an anecdote. It might have happened that Babylonians never were wrong about Venus. Moreover, the Greek origin of the names "Hesperus" and "Phosphorus" suggests that something is not right in this story. However, the anecdote helps us to introduce a problem that distresses philosophers since many years. The first one who noticed it was Gottlob Frege (1892). He rose questions that laid the foundations for a large debate in philosophy of language and the answers that he provided still exert their influence today. Let us examine the so-called Frege's puzzle[103]. Problems arise when people, later, discovered that the two names picked out the same object (presumably, Babylonians also discovered that Venus is a planet, but we can leave aside this point). The puzzle here is that language is redundant: we have more than one word to denote a single object, and we may not realize that the two words have the same referent. As Sainsbury & Tye synthesize:

> The ancient Babylonians certainly realized that Hesperus is the same as Hesperus. How could they not? The claim that Hesperus is Hesperus is trivial, a mere instance of the law of identity. But they did not realize that Hesperus is Phosphorus. Intuitively, what the Babylonians thought in thinking that Hesperus is Hesperus is not what astronomers later thought in thinking that Hesperus is Phosphorus. On this very natural view, the thought that

[103] For a classical analysis of the puzzle, see Salmon (1986).

> Hesperus is Hesperus is a different thought from the thought that Hesperus is Phosphorus. The puzzle is to say in what this difference consists. The former thought is true if and only if Hesperus is identical with Hesperus (that is, Phosphorus). The latter thought is true if and only if Hesperus is identical with Phosphorus (that is, Hesperus). So, by one natural reckoning, the two thoughts have exactly the same accuracy – or truth – conditions. What, then, accounts for their difference? (Sainsbury & Tye 2012 : 2)

At the beginning of "*Über Sinn und Bedeutung*", Frege considers the two following identity statements:

2) Hesperus is Hesperus.

3) Hesperus is Phosphorus.

Sentence (2) has the form "a=a", while sentence (3) has the form "a=b", where "a" and "b" are either names or descriptions that denote individuals. Frege assumes that a sentence like (3) is true if and only if the object *a* just is (identical to) the object *b*. However, this explanation cannot account for the whole meaning of the identity statements. In fact, the truth of "Hesperus is Hesperus" is trivial, for the sentence is a mere instance of the law of identity. On the contrary, "Hesperus is Phosphorus" is informative, as it represents a significant empirically discovery. In order to know the truth of (3), we need astronomical investigations: we must examine the world to see whether the two objects are the same.

This puzzle is related to both language and thought since, for Frege, the linguistic issue that is at the base of the oddity is supposed to be solved at the cognitive level. According to him, the difference between (2) and (3) must be explained in term of a difference in their *cognitive value*. The statement "a=a" has a cognitive value (or meaning) that differs from the cognitive value of "a=b":

> The reasons which seem to favor this are the following: a=a and a=b are obviously statements of differing cognitive value; a=a holds *a priori* and, according to Kant, is to be labeled analytic, while statements of the form a=b often contain very valuable extensions of our knowledge and cannot always be established *a priori*. The discovery that the rising sun is not new every morning, but always the same, was of very great consequence to

> astronomy. Even today the identification of a small planet or a comet is not always a matter of course. (Frege, 1892: 209)

The Millian view is hindered because it has no way to explain the difference of meaning between the two names "Hesperus" and "Phosphorus", for they both refer to the same thing. Without a difference of meaning, it is hard to explain the difference between (2) and (3). Thus, the Millian theory is forced to claim that (3), which seems informative, is trivial (since its meaning turns out to be of the form "a = a"). So, the puzzle is: how can we account for the difference in cognitive significance between "a=b" and "a=a", when they are true?

2.3 Puzzles on Propositional Attitude Reports

The second kind of puzzles concerns sentences with subordinate clauses such as propositional attitude reports. A propositional attitude is a psychological relation between a subject and a proposition. Belief, desire, intention, discovery, knowledge, etc., are all psychological relationships between subjects, on the one hand, and propositions, on the other hand. In the discussion which follows I shall only use examples with the verb "believe", but similar puzzles arise with other propositional attitude verbs as well. Again, Frege is generally credited with having identified such kind of problems, even though he did not describe the puzzle in such terms. Beliefs reports all have the same logical structure: *x* believes that *p*. When we replace the variable "*x"* with the name of a person and the variable "*p"* with a sentence that expresses the propositional object of her attitude, we get specific belief reports[104]. For instance, by replacing "*x"* with "Hammurabi" and "*p"* with "Hesperus is visible at evening", the result is the following sentence:

4) Hammurabi believes that Hesperus is visible at evening

[104] Recanati (1993: 325 – 347).

To see the problem posed by the analysis of propositional attitude reports, we need to consider a quite intuitive principle of reasoning, namely the *Principle of Identity Substitution* (also called *Substitutivity Principle*)

> **Principle of Identity Substitution**
>
> If two direct referential terms are co-referring, they are substitutable *salva veritate*, i.e., without change in truth value.[105]

The Millian theory, according to which the contribution of a referential term to the proposition expressed by the utterance in which it occurs is its reference, and nothing else, entails the Principle of Identity Substitution. But belief reports like (4) are obvious counter-examples to this principle. In fact, from (4) it does not follow:

5) Hammurabi believes that Phosphorus is visible at evening

It is perfectly plausible for Hammurabi to believe that Hesperus, and not Phosphorus, is the planet he is watching at evening, even though the names "Hesperus" and "Phosphorus", as a matter of fact, have the same referent. That is because Hammurabi does not realize that the heavenly body he calls "Hesperus" is identical with the heavenly body he calls "Phosphorus". Now, the Millian theory should predict that (4) and (5) attribute to Hammurabi a pair of beliefs that no minimal rational subject can hold, since they are contradictory: the belief that an object *x* is visible at evening and the belief that the very same *x* is not visible at evening.

Following Quine (1956)[106], philosophers distinguish between *opaque* and *transparent* contexts: a context is transparent when you can substitute any co-referential term by preserving the truth value; a context is opaque when you can not. Propositional attitude verbs,

[105] Recanati (1993: 325).

[106] Quine's distinction between opaque and transparent contexts is aimed to support his distinction between *de re* and *de dicto* belief attributions. *De dicto* belief attribution is held to be referentially opaque.

such as “believe”, generate opaque contexts that seem hard to accommodate into a Millian account[107].

The puzzle mentioned above is one of the best known arguments against Millianism. Over the years, it has been proposed in many different variations and several philosophers discussed abundantly on the matter[108]. In the following section, I shall focus my attention primarily on the versions of the puzzle elaborated by Kripke, which represent a fundamental contribution to the debate. In his series of lectures given at Princeton University in 1970, subsequently published under the title of *Naming and Necessity* (1980), Kripke argued in favor of a direct reference theory. Among other things, he claimed that the problems generating the Frege's puzzle do not concern the semantics of proper names and thus the puzzle cannot be used to disprove any semantic theory about them. Then, in his *A puzzle about belief* (1979), he presented an array of cases aimed to support his thesis[109]. The first example here discussed is based on the assumption of a very plausible principle:

> **Translation Principle (TP)**
>
> If a sentence of one language expresses a truth in that language, then any translation of it into another language also expresses a truth in that other language.

Suppose that a monolingual (French-speaking) boy, Pierre, hears about a city called “*Londres*” by reading a French book on London and, on the basis of what he reads and the pictures printed in the book, he is disposed to believe a sentence like:

6) *Londres est jolie*

[107] See Recanati (1993: 326-7).

[108] For an exhaustive presentation of the topic, see McKay, T. & Nelson (2014).

[109] As Donnellan points out in interpreting Kripke's paper: “the same paradox upon which the direct reference theory is supposed to founder, can be derived from just those principles alone, without any theory of reference being assumed at all. In other words, quite independently of any theory of reference we might adopt, there is a paradox which can be derived from principles we use in ascribing beliefs to people. And those who pose the objection to the direct reference theory must implicitly rely upon those principles. [...] The principles, of course, result in paradox only in certain cases - the very cases which are used to attack the direct reference view.”(1990: 207).

It can properly be inferred, by applying TP, that Pierre believes that London is pretty. Subsequently, Pierre moves to London where he learns the native language by exposure, without the translation of any of the words of this new language into French. So he does not realize that London is the same city he used to call "*Londres*". Since he is confined in a rather dreary part of the city and he picks up the name "London" in connection with the place he lives in, Pierre comes to believe:

7) London is not pretty

Since Pierre does not realize that the city where he lives is the same city depicted in the nice pictures he saw when he was in Paris, he has not updated his earlier belief expressed once by (6). Given that the names "*Londres*" and "London" (just as the old-fashioned "Hesperus" and "Phosphorus") have the same reference, it follows that Pierre has a set of inconsistent beliefs: he believes both that London is pretty and that London is not pretty.

It is important to note that we derive the puzzle about London/*Londres* without assuming the thesis of direct reference or relying on any substitution principle seemingly entailed by the thesis of direct reference. Kripke thus concludes that simply denying the thesis of direct reference is not sufficient for solving the original puzzle discussed by Frege. Similar problems arise every time we deal with translation between different languages: for example, a rational subject who does not know that the word "*chat*" is the French translation for "cat", may believe (8) while disbelieving (9):

8) All cats have tails

9) All *chats* have tails[110]

About these cases, Sainsbury & Tye observe:

> In the puzzle of the cat and le chat, "*chat*" translates "cat". Likewise, in the related puzzle of London and Londres, "*Londres*" translates "London". These puzzles bear some affinity

[110] This puzzle is due to Brian Loar (1987).

> to the puzzle of Hesperus and Phosphorus. They differ in that it seems more promising to make a Fregean response to Hesperus/Phosphorus cases: it seems reasonable to say that "Hesperus" and "Phosphorus" differ in meaning, or that the concept HESPERUS contributes differently to thought than does the concept PHOSPHORUS. By contrast, "cat" translates "*chat*", and "*Londres*" translates "London", so it seems hard to find a difference of linguistic meaning, or a difference of conceptual contribution.
>
> (Sainsbury & Tye 2012 : 11)[111]

This conclusion is further reinforced when we consider cases of "partial understanding" like the one presented by Burge (1979). In his example, a patient, Alfred, claims to suffer from arthritis, even though he does not know that the ill is confined to the joints and mistakenly believes that he has arthritis in his thigh. One explanation of this fact is that the linguistic community partly determines which concepts a speaker possesses, for speakers come to grasp many concepts through linguistic interaction with others. The role of the community becomes manifest in the phenomenon of deference. For example, Alfred accepts to be corrected in his use of "arthritis" by the doctor, since he thinks that the doctor is a more competent user – an expert about that term – than himself.[112] Deference to experts supports the view that competent and less competent people share the same concepts. Speakers can properly be credited with grasp of a concept, even if they do not know how to use it correctly. Only experts know which things the concept applies to. Turning back to Kripke's example, we can say that, even when Pierre uses the word "*Londres*", he has a belief about London, not about an imaginary city somewhat related to the real one, since he defers to his childhood book and to his original linguistic community.

Kripke presents another version of the puzzle that does not involve TP. This second puzzle is known as "the Paderewski puzzle". Ignace Paderewski (1860–1941) was a popular Polish pianist and composer. He gave many concerts over several decades in Europe and the United States, playing Beethoven, Chopin, and Liszt as well as some of his own works. We are told that he was a striking presence on stage, having a mop of golden red hair and a look that made women's heart flutter. But Paderewski was also a politician and spokesman for

[111] The cat/*chat* puzzle is connected to a worry raised by Benson Mates. Mates (1952) asked whether synonyms can be substituted *salva veritate* in propositional attitude contexts. If they can, then (assuming, as Mates did, that "Greek" and "Hellene" are synonyms) we could infer "James wonders whether all Greeks are Greek" from "James wonders whether all Greeks are Hellenes", but surely this is not the case. For a discussion, see Sainsbury & Tye (2012: 11).

[112] A similar point is made by Putnam's (1975) in support of the thesis of the division of linguistic labor.

Polish independence. Shortly before the First World War, he began to speak out on behalf of the cause of Polish independence and quickly became known as the most famous advocate of the Polish cause. After the War, he was the prime minister of Poland and represented his nation at the Paris Peace Conference in 1919. Once Poland had established itself as a separate state, Paderewski resigned as prime minister and returned to the concert hall, but when Poland was invaded at the beginning of the Second World War, Paderewski returned to political activities in support of his country, giving up his musical career.

Suppose now that Peter read, in one book, about a politician called "Paderewski" and, in another book, about a pianist called "Paderewski". Peter has no idea that Paderewski-the-politician and Paderewski-the-pianist are one and the same person, on the contrary, given their very different activities, he believes that they are two different people sharing a common name. Peter also believes that no politician has musical talent and that no pianist is a political figure. Knowing all the relevant facts, it certainly seems plausible to attribute to Peter both the belief that Paderewski had musical talent and the belief that Paderewski had no musical talent. But then, which beliefs have Peter? He cannot have both, it seems, for they are straightforwardly inconsistent and Peter is, by assumption, a fully rational subject.

This puzzle differs from the Fregean case, especially if we think that we are dealing with a single proper name, "Paderewski", and not with two names, like "Hesperus" and "Phosphorus". Moreover, the puzzle here does not directly concern the identity statements, but it is precisely tied to belief ascriptions:

> If you get Peter to think about politics and you ask him "Does Paderewski have musical talent?" the likely answer is "No". In many circumstances, this gives good reason to say that the subject does not believe that Paderewski has musical talent. But we've already seen that there are strong reasons to suppose that Peter does believe that Paderewski has musical talent. In the standard puzzle, the contradiction is confined to Peter's beliefs, and the puzzle is how a rational person could have such beliefs. In the strengthened puzzle, the very description of the case seems to involve a contradiction: Peter both does and does not believe that Paderewski has musical talent. (Sainsbury & Tye 2012 : 13)

Any adequate theory of propositional attitude ascriptions must account for what is going on in Kripke-style cases.

2.4 Descriptivism and Singularism

As Sainsbury & Tye observe:

> Naïve Millianism seems to mischaracterize atomic nominative concepts in two ways. It's too demanding, since it requires that every such concept has a referent, whereas intuitively there are usable nominative concepts that do not. This makes it impossible for Millians to give any natural account of empty content. It's not demanding enough, since Millian atomic nominative concepts contain no resources with which to accommodate the Fregean data. (Sainsbury & Tye 2012: 22)

Frege's answer to his own puzzle was to postulate a second level of meaning besides reference. For him, the difference between

2) Hesperus is Hesperus

3) Hesperus is Phosphorus

must be explained in terms of a difference in their "cognitive value", that is, not merely at the semantic level of reference. Names and other singular terms (like definite descriptions), in addition to reference, also have what he called a *sense* (*Sinn*, in German), which is the specific way an object is "given" to the subject:

> It is natural, now, to think of there being connected with a sign (name, combination of words, letter), besides that to which the sign refers, which may be called the referent of the sign, also what I would like to call the sense of the sign, wherein the mode of presentation is contained. [...] The referent of "evening star" would be the same as that of "morning star", but not the sense. (Frege 1892: 209)

Frege's two-dimensional semantics ascribes thus two systematic aspects of meaning, in contrast with the traditional Millian one-level semantics. The sense of an expression accounts for the cognitive significance of that expression; it is the way by which one conceives of its referent.

Many interpretations have been given about Frege's notion of sense, which he left somewhat obscure. On a standard reading, senses are understood as essentially descriptive[113]. The referent is presented with certain properties or in relation to other entities. Since the sense of an expression is supposed to determine its reference, a unique object must have the relevant properties or stand in the relevant relations to other entities. So a sense can, in principle, be made fully explicit by means of a *definite description* "the *F*". The unique object which satisfies the descriptive condition *F* is the referent.

Therefore, there is no difference in principle between a proper name like "Hesperus" and an expression like "the Evening Star": the latter may be the definite description that exemplifies the meaning of the name "Hesperus", if we consider "the Evening Star" elliptical for a description like "the first heavenly body to become visible in the evening sky". Similarly, the sense of "Phosphorus" and of "the Morning Star" may be "the last heavenly body visible in the morning sky". Distinct but co-referring expressions can introduce distinct senses, and this leads to a straightforward explanation of why "Hesperus is Hesperus" and "Hesperus is Phosphorus" express different thoughts. The pairs coincide in reference, but differ in sense. It is then easy to provide a solution to the puzzle: an identity statement of the form "a=b", like (3), is informative because it claims that one and the same planet is visible either in the evening sky and in the morning sky, whilst an identity statement of the form "a=a", like (2), is trivial because it is a mere repetition of the same definite description. By this way we can also explain why a rational subject – Hammurabi in our example – can ascribe, by believing (2) and disbelieving (3) at the same time, contradictory properties to what is, as a matter of fact, the same object: simply, Hammurabi associates different descriptions to the two terms. Finally, we must add that this picture also provides a possible explanation for empty names, like "Ulysses". The associated descriptive sense, for example the one made explicit by the definite description "the Greek hero who sailed for ten years

[113] This interpretation of Frege is resisted by some philosophers, like Evans (1982). However, it is clear that Frege, at least sometimes, uses definite descriptions as a way of describing the sense of proper names.

after the Trojan War", is one that no individual satisfies: here we have a sense without a referent.

Such an interpretation of Frege's sense directly leads to a theory known as *Descriptivism*:

> Descriptivism is the view that our mental relation to individual objects goes through properties of those objects. What is given to us are, first and foremost, properties whose worldly instantiation we are able to detect, and only indirectly objects. That is so because (according to the view) our knowledge of objects is mediated by our knowledge of their properties. Objects are given to us only qua instantiators of whatever properties we take them to have. (Recanati 2010: 141)

In a few words, Descriptivism is a semantical and epistemological thesis[114]: it claims that we can refer to singular objects of the world only by means of their properties and we can have only indirect knowledge about them. On this view, you cannot think of Hesperus just as such; the best you can do is to think of certain attributes, like "being the first star to become visible in the evening". You pick out Hesperus as the unique possessor of this property, that plays the role of a medium between your mind and the objects.[115]

The opposite view is *Singularism*. It claims that our thought is about individual objects as much as it is about properties. Objects are given to us directly, in experience, and we do not necessarily think of them as the bearers of such and such properties (even though the properties of objects are revealed to us when we encounter them in experience). As Recanati says:

> Singularism holds that our thought is about individual objects as much as it is about properties. [...] We can think of individual objects in two ways, according to Singularism. We can think of them directly, if we are acquainted with them in experience; or we can think of them indirectly, qua bearers of such and such properties. It can be maintained that

[114] It is important to underline that Descriptivism is not a metaphysical thesis. As Jackson puts it, to argue for Descriptivism "is not to advance the controversial view that objects are bundles of properties, it is to insists that we access objects via their properties" (1998: 216).

[115] At this regard, Stalnaker observes: "The descriptive strategy is to explain the capacity to refer to concrete individuals in terms of a capacity to refer to the properties and relations that are exemplified by such individuals, things that might more plausibly be thought of as internal to the mind, or at least as things that the mind could grasp from inside" (2008: 12).

> the content of a "descriptive thought" - a thought that is only indirectly about individual objects - is a general proposition, i.e. a proposition that involves only properties; but Singularism differs from Descriptivism in holding that, in addition to such thoughts, there are also singular thoughts: thoughts that are directly about individual objects, and whose content is a singular proposition - a proposition involving individual objects as well as properties. (Recanati 2010: 142)[116]

Traditionally, Singularism has been paired with one-level semantics *à la* Mill. One of the earliest proponents of this view was Bertrand Russell. He fiercely defended, against neo-Hegelian idealism of his time, the doctrine of direct realism. According to Russell, knowledge is, or rests on, a direct relation – that he calls *acquaintance* – between the mind, from one side, and objects and their properties, from the other side: "I say that I am acquainted with an object when I have a direct cognitive relation to that object, i.e. when I am directly aware of the object itself" (Russell 1910: 108)[117]. Without such a relation of acquaintance, no genuine knowledge of the external world would be possible. In another work, he writes "we have acquaintance with anything of which we are directly aware, without the intermediary of any process of inference or any knowledge of truths" (Russell 1912: 46). An alternative formulation is to say that knowledge by acquaintance occurs when the subject has an immediate awareness of some propositional truth. Knowledge by description, by contrast, is propositional knowledge that is inferential, mediated, or indirect. Over the years, despite radical changes in his doctrines, Russell kept opposing these two kinds of knowledge, using the distinction to articulate a foundationalist epistemology in which knowledge by acquaintance is the most basic kind of knowledge[118]. His stress on acquaintance and direct reference led him to reject Frege's sense/reference distinction, on the grounds that, if reference is mediated by sense, we lose the idea of direct acquaintance and succumb to Descriptivism.

Singularism *à la* Russell has a very serious limitation. In fact, even if we admit that there is knowledge based on a direct relation between the mind and its objects, we still need to specify "what exactly acquaintance amounts to, and in particular, which entities one can be acquainted with and which one cannot" (Recanati 2010: 144). Since knowledge by acquaintance does not depend on inferences or mediations, Russell employs a dubitability test

[116] For a discussion on this topic, see also Recanati (2012: 3 – 24).

[117] See also Russell (1905: 479-480; 492-493) for an earlier passing discussion of it.

[118] As he writes: "All our knowledge rests upon acquaintance for its foundation" (Russell 1912: 48).

to determine what we can know by acquaintance. The test is aimed to prevent problems such as the ones generated by the puzzles on beliefs we mentioned above[119]. But only very few entities pass the test: those entities that are given to us in such a transparent way that no identity mistake can arise. *Sense data* – the immediate objects of our perception – are the principal candidates: in fact, while it is possible to doubt whether a table exists (since we could dream, or be hallucinating, and so far) the sensory experiences of that table cannot consistently be doubted by a person who is experiencing them. Moreover, if on two occasions we were presented with a sense-datum, we could never make a mistake about whether the same sense-datum had been presented twice, or whether two sense-data had been presented. As a result, sense data can be known by acquaintance, whereas physical objects cannot. Russell also claimed that we could be directly acquainted with memory experiences, introspective experiences (awareness of our own sense data and other internal sensations), universals, and (probably) even our own self[120]. But if we are not acquainted with ordinary physical objects, we know them only "by description", via properties which these objects possess. As Recanati complains: "once you give up Frege's sense/reference distinction in favor of a monostratal semantics *à la* Russell, you are bound to embrace some form of Descriptivism" (2012: 5). According to Recanati:

> For a singularist that option is a disaster. It enables Russell to maintain the contrast between the two kinds of knowledge - direct and indirect, by acquaintance or by description - only by so drastically limiting the first kind that Russell now appears as the champion of Descriptivism. On the resulting view, almost all of our knowledge of individual objects is knowledge by description. The most typical sort of knowledge of objects by acquaintance, namely perceptual knowledge (such as the knowledge one gains of Mont Blanc when one sees the peak), now counts as knowledge by description. Defeat has not been conceded, since the idea of acquaintance remains (and acquaintance still is the foundation for all our knowledge); but defeat has taken place nonetheless. In contrast to our knowledge of the internal world, our knowledge of the external world – our knowledge of the mountains and chairs around us – is indirect, descriptive knowledge based on properties. Descriptivism rules. (Recanati 2010: 147)

[119] See ch. 2.3.

[120] See Russell (1912, ch. 5). However, for a passage in which he is less confident in the idea of being directly acquainted with the self, see Russell (1914: 81).

2.5 General Arguments against Descriptivism

Given the problem with the Singularism *à la* Russell plus the puzzles that challenge the traditional Millian semantics, Descriptivism may seem a better and more promising view. However, Descriptivism faces even more serious objections.

An obvious problem is that it seems arbitrary to couple a name with a single description, for instance "the teacher of Alexander the Great" for the name "Aristotle". Equally suitable descriptions may be "the most famous disciple of Plato" or "the last great philosopher of Ancient Greece". Moreover, at his birth, when Aristotle was so named, none of the description mentioned above were available, since he had not yet performed any action worthy of note[121]. Descriptivism needs thus to account for the variation — sometimes quite marked — of the descriptive content that people associate to the same name, a content that may vary across time and contexts for one and the same speaker. The challenge for descriptivism is to explain how people reach to communicate with each other and share a common understanding of a name despite the different descriptive content that they have in mind. As Sainsbury & Tye observe:

> It might be that you associate the word "Hesperus" with the concept THE EVENING STAR, whereas we associate it with the concept THE PLANET OCCUPYING SUCH-AND-SUCH A POSITION IN THE EVENING SKY, and some confused person associates the word with the concept THE MORNING STAR. Yet we all use the single concept HESPERUS. The descriptivist faces the challenge of explaining how all of us can, at some level of description, count as thinking of Hesperus in the same way, using the common public concept HESPERUS. The intuitive commitment to such shared public concepts emerges in ways in which we think about belief and communication. If I assert that Hesperus is visible, it won't do for you to say, "Well, I think that Hesperus is not visible,

[121] A version of descriptivism, known as cluster descriptivism, put forward by Searle (1958) and Strawson (1959; 180 ff.) is aimed to answer to this critiques. It states that a name corresponds to a definite description that is composed by a disjunction of predicates like "the teacher of Alexander" and " the most famous pupil of Plato". But this approach does not solve the problem of the cognitive significance of "Aristotle" soon after his birth. For a discussion, see Cumming (2016).

> though of course I'm not committed to disagreeing with you, since I may be using a different concept HESPERUS". On the contrary, in such cases there is a genuine disagreement, a disagreement concerning a shared thought, even if the subjects associate different information with the shared concept HESPERUS. (Sainsbury & Tye 2012 : 24)

There are other difficulties. If we consider "the teacher of Alexander the Great" as the description for "Aristotle", we note that such a description contains a proper name, which in turn needs to be interpreted as another definite description. The hope is that this description will not mention Aristotle, nor other proper names. Otherwise, we would generate a loop, which would mean that we have not specified the semantic values in question. Like many exercises of this sort, this translation has never been carried out.[122] We may assume, following Russell and Wittgenstein's suggestions, that there are "logically proper" names, like Russell's tags for sense data, to which we arrive at a certain point of our analysis. However, apart from the practical feasibility of reducing any description to logically proper names, the fact remains that descriptivism must nevertheless rely on some form of Singularism, like Russell's one.

Maybe, the most serious challenge for Descriptivism is to account for our intuition about indexicals. An indexical is a linguistic expression whose reference can shift from context to context. Paradigmatic examples of indexicals are personal pronouns ("I", "you", "he", "she"), temporal and spatial adverbs ("here", "now", "today", "yesterday") and demonstratives ("this", "that")[123]. Two speakers who utter a single sentence containing an indexical may say different things. For instance, when both Peter and John utter "I am hungry", Peter says that he, Peter, is hungry, whereas John says that he, John, is hungry. The issue was already evident to Frege:

> In this case [the sentence "I am cold"] the mere words do not contain the entire sense; we have in addition to take into account who utters it. There are many cases like this in which the spoken words have to be supplemented by the speaker's gesture and expression, and the accompanying circumstances. The word "I" simply designates a different person in the mouths of different people. It is not necessary that the person who feels cold should give utterance to the thought that he feels cold. Another person can do this by using a name to designate the one who feels cold. (Frege 1914: 134–5)[124]

[122] See Cumming (2016).

[123] The items on this list come mostly from David Kaplan's influential work on indexicals (1977).

The traditional descriptive view of indexicals would require that the reference of such expressions is fixed by some sort of adequate descriptions. But it does not seem possible to capture the meaning of an indexical by means of descriptions that do not contain indexicals in turn. As Castañeda (1999) and Perry (1993) forcefully pointed out, for any indexical *a* and non-indexical description "the F", it is always possible for a rational subject to doubt whether *a* is the F[125]. In this regard, Recanati observes:

> Indexical descriptions are descriptions that contain indexicals. Thus "the bright thing over there" contains the indexical "there". The claim that indexical modes of presentation cannot be captured by means of objective, non-indexical descriptions is the claim that, to the extent that they can be captured by descriptions, the descriptions in question will *themselves* contain indexicals. The indexical component in them is therefore ineliminable or "essential": any attempt to cash it out descriptively will produce an indexical residue, to which the same limitation applies. This shows that the indexical component in question is, at the bottom, a *non-descriptive* component; and this (of course) argues against Descriptivism. (Recanati 2012: 33)[126]

So far, every attempt of Descriptivism to account for demonstratives and indexicals resulted weak and inadequate[127]. These referential expressions seem to be irreducible to definite descriptions, even more strongly than proper names.

[124] The treatment of the personal pronoun "I" appears to be especially problematic. Frege discussed about it also in his (1918).

[125] We may also recall the famous example made by Kaplan (1977) that points at the difference between the sentence "My pants are on fire" uttered by Paul at time *t*, and the sentence "Paul's pants are on fire".

[126] A further analysis on the indexical problem may be found in Recanati (2013b).

[127] For a closer look at the debate, see Kaplan (1977); Perry (1977; 1979) and Evans (1981). For a general discussion, see Fitch & Nelson (2016).

2.6 Kripke's Arguments against Descriptivism

The attacks against Descriptivism that originate from Kripke's *Naming and Necessity* (1980) deserve special attention. His objections mostly influenced the debate over the last decades and led many philosophers of language to definitively reject the descriptivist theory. These are: first, the epistemic problem (sometimes referred to as the unwanted necessity problem); second, the modal problem (sometimes referred to as the problem of rigidity); and, third, the semantic problem (sometimes referred to as the problem of ignorance and error)[128].

For Kripke's *epistemic argument* (1980: 78, 87), let us suppose that the definite description that most aptly expresses the meaning of "Aristotle" is "the last great philosopher of antiquity". Then a sentence like (17) would be analytic, and so necessary and knowable *a priori*:

17) Aristotle was a philosopher

Such a sentence should sound as trivial as "Bachelors are unmarried" or "Squares have four sides". But it is not, since we can conceive the idea that Aristotle might never have gone into philosophy. Had things been different, Aristotle might (for instance) have died in infancy or pursued a career in the theater[129].

Kripke's *modal argument* (1980: 48-9) contends that names and definite descriptions differ in their "modal profiles"[130]. For Kripke, names are "rigid designators": they refer to the same individual in every possible world in which that individual exists. Definite descriptions, in contrast, do not appear to be rigid: the definite description "the last great philosopher of

[128] For a general discussion, see Reimer & Michaelson (2017) and Cumming (2016).

[129] Kripke argues (1980, 68ff) that even the sentence "Aristotle is the individual called 'Aristotle' "—supplied by the metalinguistic descriptivist theory—is not knowable *a priori*.

[130] See also Marcus' proof of the necessity of identity statements involving names (1947).

antiquity" might well refer to Plato in a world where Aristotle died in infancy. But it seems intuitively implausible to suppose that the name "Aristotle" — as we use that name in the actual world — could be used to pick out anyone other than its referent in the actual world, namely Aristotle. Kripke supports these intuitions about proper names by appealing to "might have" modal sentences (taken in the "ontic" or "metaphysical" meaning rather than in the epistemic one). For instance, while the first sentence sounds true on this reading, the second sounds false:

18) Aristotle might not have been the teacher of Alexander

19) Aristotle might not have been Aristotle

However, if we adopted Descriptivism, we would be forced to say that sentence (18), albeit intuitively true, is false. This suggests that names are semantically different from descriptions, which in turn suggests that the mechanism by which a name refers cannot be equated with the one used for definite descriptions.

The *semantic argument* is due either to Donnellan (1966, 1970) and Kripke (1980, 80ff), and it is related to the externalist arguments of Putnam (1975) and Burge (1979). We can identify two distinct versions of the argument. The first is the *problem of error*. It consists in pointing out that people may associate inadequate or inaccurate information with proper names. For instance, let us suppose that someone knows about Einstein only that he was a physicist. In this case, the cognitive value of the name "Einstein" corresponds to either an indefinite description ("a physicist") or an improper definite description ("the physicist"). Then, our speaker would fail to refer via his use of the name "Einstein", since the associated descriptive content is not sufficient to pick out Einstein from the mass of countless other physicists[131]. The second is the *problem of ignorance*. This argument claims that even when people associate well-formed descriptions with proper names, those descriptions may be wrong and thus not to determine the correct reference. For instance, suppose that the cognitive significance carried by the name "Gödel" is exemplified by the description "the discoverer of the Incompleteness Theorem" (under the assumption that this is the only thing that people know about Gödel), yet it turns out that Gödel did not really prove it, but he stoled the

[131] Note that even the metalinguistic information "the physicist called 'Einstein' " is still insufficient if there is more than one so-named physicist. See Burge (1973).

theorem from his friend Schmidt. The point is that, in such a possible world, the descriptive meaning associated with "Gödel" picks out an individual that is not Gödel: Schmidt becomes the referent of the name.

The argument of ignorance can also be applied to definite descriptions, as it happens in the famous example made by Donnellan:

> One is at a party and, seeing an interesting-looking person holding a Martini glass, one asks, "Who is the man drinking a Martini?". If it should turn out that there is only water in the glass, one has nevertheless asked a question about a particular person, a question that is possible for someone to answer. (Donnellan 1966: 48)

For this reason, we can distinguish, borrowing the terminology from Bach (1987), between descriptions used "*referentially*" as opposed to "*attributively*" (or "*satisfactionally*")[132]. A referential use presupposes a certain relation (for instance a perceptual relation) between an object and a speaker who intends to refer to the object by means of a description she believes the object satisfies. As Bach claims:

> Since the object of a descriptive thought is determined *satisfactionally*, the fact that the thought is of that object does not require any connection between thought and object. However, the object of a *de re* thought is determined *relationally*. For something to be the object of a *de re* thought, it must stand in a certain relation to that very thought.
>
> (Bach 1987: 12)[133]

Such considerations led a number of philosophers, in the late sixties and early seventies, to argue in favor of an "externalist" approach to reference determination. According to Kripke and Donnellan[134] (and, anticipating them, Geach 1969), what determines the reference of a name (or of a definite description referentially used) is not a description in the head of the speaker, but historical facts about that use and the communicative chain to which it belongs. Kripke (1980: 91) suggests that the reference of a name is established by a

[132] See Recanati (2012: 18).

[133] See also Bach (1986: 188-9).

[134] See also Devitt (1981) for further development of the theory.

dubbing ceremony (or "baptism") in which the dubee is indicated by a demonstrative or a uniquely referring description. All uses of the name that derive from this source (uses deriving from the baptism itself, or acquired from someone who was present at the baptism, or from someone who acquired it from someone who was present at the baptism, etc.) refer to the original dubee, even if the speaker associates the name with a description that is untrue of that dubee. Kripke includes the following caveat in his account of the reference-passing links through a causal-historical chain:

> When the name is "passed from link to link", the receiver of the name must, I think, intend when he learns it to use it with the same reference as the man from whom he heard it. If I hear the name "Napoleon" and decide it would be a nice name for my pet aardvark, I do not satisfy this condition. (Kripke 1980: 96)

Although the causal theory provides a plausible account of nominal reference, its advocates still need to supplement their theory of reference with a theory of meaning – a theory that accounts for the fact that proper names appear to have some sort of meaning or cognitive content.[135] On its own, the causal theory of reference does not provide any answers to the questions of cognitive significance that so bothered description theorists like Frege and Russell.

2.7 Non descriptive modes of presentation

The descriptivist interpretation of Fregean modes of presentation is not the only one available. Even if it is true that Frege had clear descriptive tendencies, his two-level semantics does not necessarily entail the indirectness of our thought. As Evans (1981) fiercely remarked, it is a quite unwarranted assumption the view that the Fregean sense of any singular term must be expressed by means of a definite description. To explain this point, we need to make clear

[135] See Reimer & Michaelson (2017).

which role a sense is supposed to play and, consequently, which is the key criterion for individuating the best candidates for the role of sense. It is generally recognized that senses must obey what Schiffer calls *Frege's Constraint*:

> Necessarily, if *m* is a mode of presentation under which a minimally rational person *x* believes a thing *y* to be F, then it is not the case that *x* also believes *y* not to be F under m. In other words, if *x* believes *y* to be F and also believes *y* not to be F, then there are distinct modes of presentation *m* and *m'* such that *x* believes *y* to be F under *m* and disbelieves *y* to be F under *m'*. Let us call this *Frege's Constraint*: it is a constraint which any candidate must satisfy if it is to qualify as a mode of presentation. [...] In effect, Frege's Constraint provides *the* motivation for introduction of modes of presentation (Schiffer 1978: 180)[136]

In few words, senses must allow us to account for situations like the ones presented in ch. 2.3, in which Hammurabi believes that Hesperus, and not Phosphorus, is visible in the evening sky. But to reach such a target, it is sufficient to admit that "one must not merely think to the reference that is the reference, but that one must, in so thinking, think of the reference *in a particular way*" (Evans 1981: 294). Frege's Constraint does not force us to adopt a substantial theory about what senses are. On this issue, Evans states:

> In particular, it is not necessary [...] to suppose that ways of thinking of objects can always be given by giving some definite description uniquely true of the object, or to make any other supposition which would lead to "existential-independent" senses. It is not necessary, because it is not plausible to suggest that the only kind of account of what makes a subject's thought about an object which is capable of making (P) true is one which relies upon the subject possessing a unique description of the object. (Evans 1981: 301-2)

We can make room for *existence-dependent* Fregean senses, i.e. senses that are *acquaintance-based,* modes of presentation that directly point at the object, in the same way as Russellian singular terms do. Modes of presentation are now construed as "ways the object is given to the subject" (Recanati 2012: 12): an object can be given directly through acquaintance, or indirectly via properties which it uniquely instantiates.

[136] Evans provides an analogous criterion (1981: 301).

Note that we do not need to give up descriptive modes of presentation. It simply means that we can refer to an object either directly or by means of a definite description. The first case is presumably involved in a demonstrative sentence like

20) That mountain [pointing at the Mont Blanc] is 4.000 meters high

In uttering (20), the subject must be in a proper relation of acquaintance to the object the thought is about. It is not possible to understand (20) if we do not share a similar kind of relation, for example if we cannot see which mountain the subject is pointing at. On the contrary, when uttering

21) The tallest mountain in Europe is 4.000 meters high

no acquaintance is presupposed. Everyone can grasp the meaning of (21), even children that are listening to the geography teacher's lesson and have never seen the Mont Blanc.

It is important to underline that we are here proposing a notion of reference by acquaintance that is different from the one suggested by Russell. As Recanati observes:

> The idea of "directness" turns out to be ambiguous. "Direct reference" can mean that the only meaning or content of a representation is its reference, to the exclusion of any sense or mode of presentation, as Russell's one level semantics; or it can mean, as in singularist framework, that the subject is directly acquainted with the object in the experience and does not think of it descriptively as the instantiator of such and such properties. The two ideas are clearly independent, and it was a mistake on Russell's part to argue from Singularism to the rejection of Frege's two-level approach. (Recanati 2012: 13)

Contrary to the idea that Singularism must be paired with a Millian one-level approach to semantics, modes of presentation are compatible with reference through acquaintance, in a way that allows Singularism to overcome the limits that affect Russell's account.

2.8 Mental files

In the paragraph above, we explored the reasons that led some philosophers to admit the existence of non-descriptive, or object-dependent, modes of presentation. But what exactly are non-descriptive modes of presentation? A promising suggestion is that they can be described in terms of mental files.

I shall base my analysis on François Recanati's works (1993, 2010, 2012, 2016), which represent an important contribution to the debate. As for the background, the notion of mental files benefits from a few decades of history. A first detailed account is due to Strawson in his discussion on identity statements (1974: 54-56), even if a forerunner notion may be found in Grice, which rather incidentally used the term "dossier" in connection with referential descriptions (1969: 140). However, the most influential source is probably due to Perry, who introduced the term "mental file" for the first time to account for the phenomenon of continued belief (1980)[137] and went on developing the idea in his further works (2001a, 2002). Many other authors have subsequently exploited Perry's notion. Other important recent contributions to the theory of mental files are the works of Sainsbury & Tye (2012) and Tim Crane (2013). They use the term "concept" or "singular concept" instead of mental file, but the notions are comparable[138].

According to Recanati, the notion of mental file translates, at the cognitive level, the Fregean idea of sense of a proper name or a singular term.[139] It is a cognitive structure that we use to create a mental representation of particular objects and, speaking metaphorically, we can imagine a mental file as a box where we put information, under the form of a list of predicates, that we take to be about an object of the outside world. For example, since in my mind I associate to the individual called "Aristotle" properties like "the teacher of Alexander the Great", "the most famous disciple of Plato", "the last great philosopher of Ancient

[137] Donnellan (1970, 1974) is credited by Perry (2000: 70-74) for the basic inspiration.

[138] See Crane (2013: 138- 167).

[139] For a critical discussion on this point, see Pagin (2013).

Greece", or even more general and vague notions like "a Greek philosopher", "a wise man of the past", etc., it means that I have a mental file about Aristotle in which there are stored such and such pieces of information. But we must distinguish between the container – i.e., the mental file – and the content – the information stored in the file. In this regard, Crane clearly states:

> We can distinguish between the representation itself, and the body of general information associated by the speaker with the representation. [...] We should not think of the information in the file as the *meaning* of the name or other expressions which we use to express the thought in question. The meaning of a term is something which is given by a correct semantic account of that part of the language. What a term means in a public language may be something which goes beyond any information a thinker may have about the referent of the term, and the information a thinker has may be far richer than the meaning. (Crane 2013: 158)

The distinction is crucial because it makes possible for files to take the role of Fregean senses, without falling into Descriptivism[140]. In fact, the reference of a file is determined in a non-descriptive way through the specific relation that the subject entertains with the object of her mental representation:

> A non-descriptive mode of presentation, I claim, is nothing but a mental file. Mental files are based on [...] "acquaintance relations". [...] Different types of file correspond to different types of relation. The role of the files is to store information about the objects we bear these acquaintance relations to. So mental files are "about objects": like singular terms in the language, they refer, or are supposed to refer. They are, indeed, the mental counterparts of singular terms. What they refer to is not determined by properties which the subject takes the referent to have (i.e. by information - or misinformation - in the file), but through the relations on which the files are based. The reference is the entity we are acquainted with (in the appropriate way), not the entity that best "fits" information in the file. (Recanati 2012: 34-35)

Thus, mental files rest on acquaintance relations with the objects in the environment. The reference is determined, using Bach's terminology (1987), relationally and not

[140] About the distinction between mental files and their content, and for an analysis of its theoretical role, see also Recanati (2012: 38- 41).

satisfactionally: a mental file refers to the object the subject is acquainted with, whether or not such an object fits the subject's representation of it.

But what are these relations we are talking about? The characteristic feature of the relations on which mental files are based is to be *epistemically rewarding* (ER): i.e, the relations enable the subject to capture information about the object we are related with:

> In his cognitive life the subject encounters various objects to which he stands in various contextual relations. Some of these relations – the acquaintance relations – are epistemically rewarding in that they enable the subject to gain information from the object. For example, by holding an object in my hand, I can get information about its weight. By looking at it I can get information about its visual appearance. The role of a mental file based on a certain acquaintance relation is to store information acquired in virtue of that relation. (Recanati 2010: 156-7)[141]

There are three main types of ER relations, and to each specific ER relation corresponds a different type of mental file. The paradigmatic case, such as the one presented above, is *perceptual acquaintance*. We have perceptual acquaintance when we perceive an object directly with our senses, by seeing, hearing, or touching it. It is the sort of relation to objects which makes the perceptual flow of information possible. Note that we are able to have a thought and a mental representation of an object that we perceive even if we may not be capable of providing an adequate description of the object. While walking in the middle of a forest, I can stop and admire a massive tree in front of me, without having in mind a description of this tree that uniquely specifies the referent. I may have an indefinite description ("a giant tree"), or a definite description that, in this context, fails to identify a single object ("the tree that is as high as a house"). I may even generate a wrong description, because I mistakenly take a beech to be an elm, like in Putnam's example (1975). The object the thought is about is the object the perception is about: the reference is not determined by what the subject believes on it. As Campbell puts it:

> Your visual system is managing to bind together information from a single thing, and you are consequently able to attend consciously to it, even though you have not managed to apply the right sortal concept to it. (Campbell 2006: 205)

[141] See also Recanati (2012: 34-38) and (2014).

Of course, even if I am not able to conceptually articulate what my thought is about, it is always possible to find a reference-fixing description of the sort "what I am now seeing" or "what I am now touching". But these are indexical descriptions, that face the problems mentioned in ch. 2.6, i.e. the indexical element cannot be eliminated. Moreover, this move seems to be *ad hoc* because, as Recanati observes, such a kind of descriptions presuppose, on the part of the subject, reflective abilities which appear superfluous: "the subject need not reflect on her perceptual relation to objects in order to have thoughts about the objects she perceives; nor does she have to reflect on her perceptual relation to objects to be in a position to think of the object in different ways, corresponding to the various ways she perceives it" (2012 : 30)[142].

A perceptual relation typically generates short-terms files that exist only as long as acquaintance lasts. They may be called perceptual files (perceptual "buffers", following Perry 2001a) or demonstrative files, since it is a mental representation of this kind that we have in mind when we use demonstratives, for instance by saying "That's a beech, not an elm!". As long as we are in the right type of perceptual contact with the object, we can think of it demonstratively. When the relation is broken, the temporary files based on it disappear. The information stored in the files, however, is not lost but transferred into another kind of files[143].

The second case of ER relation takes place when acquaintance is indirect or mediated, established through communicative chains, as it is for people that we do not know directly or that lived in the past. I had not the opportunity to meet Socrates, as he was born many centuries before me, but I know that that person existed and did such and such things during his life because some other people were directly acquainted with him at that time, and they taught to the subsequent generation who Socrates was. That generation has done the same with the next one, and so on through the centuries. As Geach says:

> For the use of a word as a proper name there must in the first instance be someone acquainted with the object named. But language is an institution, a tradition; and the use of a given name for a given object, like other features of language, can be handed on from one generation to another; the acquaintance required for the use of a proper name may be mediate, not immediate. Plato knew Socrates, and Aristotle knew Plato, and Theophrastus

[142] See also Recanati (2014).

[143] See Recanati (2010: 157).

> knew Aristotle, and so on in apostolic succession down to our own times; that is why we can legitimately use "Socrates" as a name the way we do. (Geach 1972: 155)

It is the existence of this unbroken tradition of testimony that enables me to acquire information about Socrates. Such informational channel is *de facto* an ER relation putting in contact the subject with the individual at the origin of the chain. We have a mediated acquaintance every time the learning presupposes a sequence of witnesses, no matter whether long or short, whether oral or written. I suppose that we can admit even pictorial testimony, for example when you are looking at the map to determine the route to follow in the city that you are going to visit in the summertime. In all cases, there is a set of community-mediated testimonial relations to objects. This idea takes inspiration by what Lewis says:

> [The notion of acquaintance can be generalized] in virtue of the analogy between relations of perceptual acquaintance and others, more tenuous, relations of epistemic rapport. There are relations that someone bears to me when I get a letter from him, or I watch the swerving of a car he is driving, or I read his biography, or I hear him mentioned by name, or I investigate the clues he has left at the scene of his crime. In each case there are causal chains from him to me of a sort which would permit a flow of information. Perhaps I do get accurate information; perhaps I get misinformation, but still the channel is there. I call such relations as these *relations of acquaintance*. (Lewis 1999: 380-81)

Note that it is not necessary and, as a matter of fact, it is very unlikely, that the subject is aware of the communicative chain that relates her to the object. As Geach pointed out, the existence of a communicative chain is sufficient to enable the speaker to successfully refer to a particular object. It is not required for such a communicative chain to be represented, any more than the perceptual relation to the referent has to be represented in order for the subject to successfully refer to an object she is acquainted with:

> It is not our knowledge of this chain that validates our use, but the existence of such a chain; just as according to Catholic doctrine a man is a true bishop if there is in fact a chain of consecrations going back to the Apostles, not if we know that there is. (Geach 1972: 155)[144]

[144] A similar remark can be found in Kripke (1980: 93).

The alleged challenge by descriptivists thus does not hold: in fact, even if it is possible to provide a metalinguistic description like "the person called 'Socrates' ", a description that is satisfied by whoever stands at the other end of the communicative chain, a description of this kind appears unnecessary and superfluous in this context[145].

As opposed to demonstrative files, the ones based on mediated acquaintance are typically long-lasting files, because the ER relation established through a communicative chain lasts longer than transient perceptual acquaintance.

Lastly, we have contextual relations to objects as those exploited by indexicals like personal pronouns "I", "you" or temporal and spacial adverbs "now", "here". Like demonstratives, indexicals present their reference in a perspectival way and their meaning cannot be captured by means of definite descriptions that do not contain in turn some indexical residue. As a matter of fact, we could consider, following Kaplan and Perry, demonstratives as a sub-class of indexical expressions, and so reduce the first class of files to the latter. Indexicals are short-term relations and so are files generated by them, but not always. A particular case is, in fact, the SELF file. According to Perry (2000, 2002), the concept of self is a type of mental file that is based upon a special relation which every individual bears, permanently, to himself or herself, namely identity. With regard to this point, Recanati observes:

> In virtue of *being* a certain individual, I am in a position to gain information concerning that individual in all sorts of ways in which I can gain information about no one else, for example, through proprioception and kinaesthesis. The mental file SELF serves as repository for information gained in this way. In contrast, the files associated with other indexicals ("here", "now" ...) are based on short-lived ER relations to the place we are in, or to the current time, which relations enable the subject to know (by using his senses) what is going on at the place or time in question. They are similar to perceptual buffers, which is to be expected given the link between indexicality and perception. (Recanati 2012: 36)

To conclude, we can say that mental files allow us to form a mental representation of singular objects of the world we are acquainted with. Regardless the specific type of ER relation a file is based on, whether perceptual, indexical or mediated by a communicative chain, the role of a mental file is to store information acquired in virtue of such a relation. The information may

[145] For this point, see Recanati (2012: 32-33).

be wrong or inaccurate, but this does not prevent the mechanism of reference determination, that is "relational" and not "satisfactional". What determines the referent is not the content of the file – the list of predicates that express the properties we associate to the object – but the relevant relation to the object. "The file corresponds to an information channel, and the reference is the object from which the information derives, whether that information is genuine information or misinformation" (Recanati 2012: 38)[146].

2.9 How to apply mental files

Let us go back to the Frege's puzzle about Hesperus and Phosphorus. How is the mental file's framework useful to account for cases like this? Since mental files play the role of modes of presentation, they work in the same way as Frege's senses. A subject, in our example the Babylonian king Hammurabi, may have in his mind two separate files, because he is acquainted with the same object through two distinct ER relations. He sees Venus in the morning, so there is in a perceptual relation justifying the creation of the file PHOSPHORUS. Then, when he observes the sky at night, a new perceptual relation arises between the king and the planet, thus a new file, HESPERUS, is generated. Hammurabi employs either the file PHOSPHORUS or the file HESPERUS to refer to Venus. The two files remain separate until Hammurabi discovers that the heavenly body he used to call "Hesperus" and the heavenly body he used to call "Phosphorus" refer to one and the same planet Venus. Hence,

3) Hesperus is Phosphorus

is informative in virtue of the distinctness of the files associated with the two terms of the identity statement. The cognitive value of (3), the idea of discovery it expresses, derives from the fact that we can now connect the files and join the pieces of information stored in each of

[146] For a deeper analysis on the mechanism of reference of mental files, see Recanati (2013b).

them. In fact, a file collects information about an object that derives from the specific ER relation the file is based on. The role of the file is precisely to treat all information as concerning one and the same object and to license integration and inferential exploitation of the information in question. But when our knowledge about an object is split between two separate files, we are not able to exploit it as a single set of data. "To say that there are two distinct mental files is to say that information in one file is insulated from information in the other file" (Recanati 2012: 42). Recanati, following Perry[147], calls *linking* the specific operation that licenses the integration of information encapsulated in separate files:

> When two files are linked, information can flow freely from one file to the other, so informational integration/exploitation becomes possible. [...] From a cognitive point of view, linking is a quite fundamental operation. It is involved, for example, in the phenomenon of recognition (which involves linking a perceptual file and a file based on memory) – or at least, in some forms of recognition. It is that operation which I think accounts for the cognitive effect of accepting an identity statement. To accept the identity "A is B" is to link the two files corresponding to the terms on each side of the equal sign. (Recanati 2012: 43 – 44)

As long as the single files remain separate, I cannot integrate the information within them, but when I link them, I can join the pieces of information that each of them contains. So, if I know that Hesperus is a heavenly body visible at evening, that Phosphorus is a heavenly body visible in the morning, and that Hesperus is Phosphorus, I can conclude that there is a heavenly body that is visible both at evening and in the morning. Consider the difference between (3) and

2) Hesperus is Hesperus

[147] The idea of links that connect different mental representations is already present in Strawson (1974). However Recanati, referring to Perry (2002), distinguishes between the notion of *linking* - a connection of files that remain distinct – from *merging* – a fusion of two files into one. In fact: "Two linked files may end up being merged, after some time (especially as new information accumulates). [...] Linking is less risky, as it can easily be undone. So merge is an option for dealing with an identity, but it should not be automatic" (Recanati 2014: 475).

In this second case, there is no link, simply because there is a single file at work that is exercised twice. The statement is a mere assertion of self-identity: we do not get any new information, nor we acquire the possibility to integrate and exploit two pieces of data that were previously isolated.

Let us consider now how mental files can be spontaneously employed to account for the second puzzle we mentioned, the puzzle related to belief ascriptions. The two files HESPERUS and PHOSPHORUS, as long as they remain separate, contain different lists of predicates, for example file HESPERUS contains the predicate "being visible in the morning sky", but not the predicate "being visible in the evening sky", that is stored in the file PHOSPHORUS. So Hammurabi, before realizing that both files refer to the same planet, may have different and contradictory beliefs about Venus, without this fact undermines his status of rational subject. Therefore,

4) Hammurabi believes that Hesperus is visible at evening

states that Hammurabi thinks of Venus, under the mode of presentation represented by the mental file HESPERUS, that it is visible at evening. But Hammurabi cannot entertain the thought that Venus, under the mode of presentation PHOSPHORUS, is visible at evening, because such information is not contained in that sort of file. This explains why

5) Hammurabi believes that Phosphorus is visible at evening

is false. Mental files mediate between us and the objects and are responsible for the way we cluster information in our mind. When we deal with a single file, data are always at our disposal for inferential exploitation. We can express the same idea by saying that senses are *transparent*: under a single mode of presentation, an object is given to us in such a transparent manner that no identity mistake can arise. That is, a single mental file cannot contain two contradictory information. For Hammurabi, to believe, by deploying one and the same file, let us say HESPERUS file, that Venus is both visible at evening and not visible at evening, would suffice to make him an irrational believer. As Recanati observes:

> That is the epistemic transparency issue. Reference, as we all know, is not epistemically transparent. The subject may not realize that two terms refer to the same object, or that they refer to distinct objects. In both cases the subject may be deluded. What about sense? It seems that, in contrast to reference, sense (mode of presentation) must be transparent. If modes of presentation themselves are not transparent, there is no reason to move from pure referential talk to mode of presentation talk in the explanation of rational behavior. The appeal to senses (as opposed to sheer reference) in psychological explanation presupposes the transparency of sense as opposed to the non-transparency of reference. [...] A rational subject must be capable of reflecting critically upon his or her own thoughts; that sort of reflexive control over one's thoughts is possible only if they are transparently accessible. Transparency follows from rationality. (Recanati 2012: 116 - 117)[148]

So far I have discussed the puzzles of identity statements and attitude ascriptions basing the analysis on the example of Hesperus and Phosphorus. It is easy to explain why Hammurabi has two distinct files referring to Venus: he looked at the planet in different occasions, i.e. in two different temporal ranges. However, we do not need to think that this sort of cases takes place only with mental files generated by perceptual relation. Analogous situations are possible also when we deal with other ER relations. I conclude the paragraph by focusing on the acquaintance derived by a community-mediated testimonial chain. Suppose that a student heard by his teacher of Latin that Cicero was a Roman philosopher. Then, suppose that, by attending a class of Ancient History, it is told to him that Tully was not a philosopher. Of course, Cicero is Tully, but our student is not aware of this fact, because he came to know about Cicero/Tully through two separate communicative chains: he thus created two separate mental files. For one and the same individual *x*, he may accept the claim that *x* was a philosopher, when he uses the file CICERO, and reject the claim that *x* was a philosopher, when he uses the file TULLY. Our student believes that the two names "Cicero" and "Tully" refer to two distinct people and that two distinct claims are made.

There is not a substantial difference between the Cicero/Tully case and the Paderewski's puzzle. It does not matter that, here, we have only one name. The subject, Peter in Kripke's example, is acquainted with Paderewski through two different communicative chains, so he has two mental files: PADEREWSKI-THE-PIANIST and PADEREWSKI-THE-

[148] For a detailed analysis of the transparency issue, see Recanati (2012, 115 – 143). About that topic, see also Boghossian (1994) and Schroeter (2007).

POLITICIAN. For Peter, it is a simple case of homonymy. The same analysis applies, even more easily, to the London/*Londres* puzzle. The latter example is also interesting because it is generated by a mix of heterogeneous ER relations: Pierre has firstly established a mediated relation with London by reading a French book and looking at its pictures, then a direct acquaintance relation comes into existence after Pierre's moving to the city.

As a final remark, I wish to emphasize once again the importance of ER relations. As already said, an ER relation is an informational channel putting a subject in contact with a particular object of the world, allowing the acquisition of new knowledge. It is this channel that identifies a mental file, not its content, i.e., the set of data stored in it. The perplexity expressed by Kit Fine is therefore easy to address:

> The main problem with the Fregean position is to say, in particular cases, what the difference in the meaning or sense of the names might plausibly be taken to be. Although there appear to be good theoretical reasons for thinking that there *must* be a difference, it seems hard to say in particular cases what it is. For as Kripke (1980) has pointed out, it seems possible for a speaker, or for speakers, to associate the same beliefs or information with two names, such as “Cicero” and “Tully”. And if the information or beliefs are the same, then how can the sense be different? (Fine 2007: 35)

I may have two distinct files, CICERO and TULLY, containing the same information, for instance the predicate “a famous Roman orator”, but until they remain separate, I have two distinct modes of presentation to think about such an individual. As Recanati says: “what plays the role of sense is not the information in the file, but the file itself. If there are two distinct files [...] then there are two distinct senses, even if the information in the two files is the same.” (2012: 40-41).

2.10 Hierarchy of files

Since the function of a file is to exploit a given ER relation, it follows that many files are ephemeral. As we have already said, a file exists as long as the subject stands in the appropriate acquaintance relation to the entity. However, most acquaintance relations, especially perceptual and indexical ones, have a short-lived term. Usually, I am not looking at a tree for more than a few seconds or minutes; then, I shift my attention elsewhere. When I am no longer in a position to perceive the tree or to focus my attention on it, the file expires: I cannot continue thinking of the object I was seeing as "this tree". Similarly, I think about the room where I currently stay as "here", since I am in such contextual relation to the space around me. But I can no longer think of a place as "here" if I no longer occupy that place. Of course, "I can still think HERE-thoughts, but the HERE-modes of presentation occurring in those thoughts will be modes of presentation of different places, hence different modes of presentation (though modes of presentation of the same *type* as my present HERE-mode of presentation)" (Recanati, 2012: 61). The only exceptions are represented by files based on long-term relations of testimony-mediated acquaintance and the file SELF, based on a relation that everyone bears to himself or herself, in a permanent way. This picture appears to be quite limited and need to be extended:

> What happens when the contextual relation to the object ceases to hold? As I have just pointed out, I can no longer think of a place as HERE if I no longer occupy that place. And I cannot think demonstratively of an object which I can no longer perceive. In both cases, however, *another* mode of presentation – another file, based on another relation to the object – becomes available as a substitute. For example, when a demonstrative mode of presentation comes out of existence because the demonstrative relation on which it is based no longer holds, another relation comes to hold, in virtue of which I *remember* the object. On that relation another mode of presentation is based, distinct from but closely related to the original demonstrative mode of presentation.
>
> (Recanati 2012: 62)

When a temporary file disappears, together with the contextual relation that generated it, the information contained in the file is not lost. It is simply transferred into a more durable file[149]. In fact, there continues to be a relation between the subject and the object, the difference is that now the object is remembered, no more perceived. Memory is, again, an ER relation, since by remembering an object the subject recalls information to be used for his present purposes[150]:

> Just as demonstrative modes of presentations are based on demonstrative relations in virtue of which one can perceive the object, memory demonstratives are based on certain relations in virtue of which one can remember the object. Through our memories of the object, we can focus our attention on it even after the perceptual encounter has ended. So we can say that the demonstrative THAT MAN [WHOM I SEE] is converted into a memory demonstrative THAT MAN [WHOM I SAW]. (Likewise, HERE can be converted into THERE, and NOW into THEN). (Recanati 2012: 62)

This kind of considerations puts the base for the foundation of a hierarchy of files[151]. It is a hierarchy because, at the bottom, we set files more closely tied to a specific ER relation, whereas, at the top, this dependence on ER relations loosens. Moreover, as we shall see, files at each level derive from the files at the previous one.

At the first step, we have the so-called *proto-files*: files that only host information gained in virtue of a specific ER relation to the referent. Proto-files are short-lived: they exist as long as the ER relations on which they are based exist. These files are very limited cognitive tools: even if we may consider them as modes of presentation, proto-files cannot play the cognitive role of individual concepts – i.e. to be mental particulars that serve to represent in thought individual objects – because they fail to satisfy a fundamental requirement, clearly formulated by Evans:

---

[149] As Recanati told me in a private conversation, we should imagine a cluster of information lasting over time, whereas the mental file is the shape we give to this cluster from time to time, according to the ER relation that is actually at state.

[150] Again, I appreciated this point thanks to a conversation with Recanati.

[151] This idea has been further developed in Recanati (2016).

> If a subject can be credited with the thought that *a* is F, then he must have the conceptual resources for entertaining the thought that *a* is G, for every property of being G of which he has a conception. This is the condition that I call "The Generality Constraint".
>
> (Evans 1982: 104)

The Generality Constraint establishes that, if we possess the concept of some object, we must be able to apply it to whatsoever predicate we attribute to the referent. Translated in the mental files framework, this means that we should be able to put in the file any kind of predicate. Clearly, that is not possible with proto-files, since they contain only the information gained in the special way that goes with a specific ER relation. It is for this reason that Recanati is reluctant to consider them as files *simpliciter*:

> Take the proto-file SELF*: it can only host information gained from inside, through e.g. proprioception or introspection. Now [...] there is much information about myself that I cannot gain in this way. My date of birth is something I learn through communication, in the same way I learn my parents' birthdate. In virtue of the Generality Constraint, it should be possible for that information to go into my SELF file, and that is the crucial difference between the SELF file and the (non-conceptual) proto-file SELF* from which it originates.
> (Recanati 2012: 65)[152]

From proto-files derive standard *conceptual files*. We can split the content of a file in two parts: information that is obtained from the relevant ER relation, and information that is not obtained in this way, but that the subject takes to concern the same individual the file refers to. Thus, we should conceive a mental file as an expansion of a proto-file. The proto-file remains as the nucleus to which other pieces of information are attached[153]. However, the mental file so formed is still based on one kind of ER relation: the file THAT-MAN is a perceptual demonstrative file, even if I know from a picture previously seen that the person in front of me is Recanati; the file YOU is an indexical file, regardless of how much information, gained from other sources, I know about the person I am speaking with; the file THAT-ELM, referred to the tree that I saw yesterday during my walk, is a memory demonstrative file. Except for

[152] See also Recanati (2010: 157).

[153] See Recanati (2012: 64 – 67).

the last one, all of them still are temporary files, which are suppressed when the relevant relation of acquaintance despires.

We may now focus on *stable files*, the last step of this tri-fold hierarchy. A special case of stable files is the SELF files. The permanence of the files rests on the fact that it is based on the enduring relation of identity with oneself that everyone entertains forever and necessarily. Then, there are what Recanati calls *recognitional files*, a further development of memory demonstrative files:

> Recognitional files are based on a relation to the object which I call "familiarity". An object is familiar to the subject just in case multiple exposure to that object has created and maintained in the subject a disposition to recognize that object. (Recanati 2012: 71)[154]

These files still depend on their existence upon a contextual relation to the object, namely the relation of familiarity, but such a relation presupposes the subject's disposition to recognize the object, which disposition transcends particular encounters with the object. So recognitional files are stable, in contrast to ordinary demonstrative files.

Finally, there is a last kind of enduring files that deserves our interest. So far we have talked about files that are based on a single contextual relation that enables the subject to acquire information. But there are cases, quite easy to observe as a matter of fact, that challenge this picture:

> I can recognize Obama, but my OBAMA concept does not reduce to a recognitional concept. I would not lose my concept if I lost my capacity to recognize Obama. The relevant concepts seem to be detachable from any specific informational relation to the referent. (Recanati 2012: 72)

According to Recanati, we must make room, in addition to files that are very closely tied to a specific contextual relation, to files that are more abstract, in the sense that overcome a single ER relation. On the one hand, there are the first-order files, or egocentric concepts[155], on the

[154] See also Recanati (2006a).

[155] See Recanati (1993, 119 – 132). See also Recanati (2010, 157-158) and (2012: 72 – 75).

other, there are second-order files, or encyclopedia entries. These latter *encyclopedic files*[156] are based on a more general-purpose tracking relation: it is a higher-order relation which holds whenever we stand in some ER relation to an object:

> An encyclopedic file may exploit a number of ER relations to the reference of the file, in an opportunistic manner, instead of being based on a single one. Any relation will do, provided it preserves the link to the object. In this case, what determines the reference of the file is the overarching tracking relation: the relation between the file and the object it has been created to track (however it is tracked). (Recanati 2012: 73)

My mental file OBAMA is not tied to a particular way of gaining information, i.e. a specific ER relation, but contains all the information that I can get about this person, however it is gained. Not by chance, encyclopedic files are typically associated with proper names or, to say better, they may be considered as the mental counterparts of proper names. Contrary to first-order files, the content of encyclopedic files is preeminently descriptive[157], even if the reference continue to be determined by means of an (higher-order) ER relation. So they still are non-descriptive modes of presentation[158].

[156] Encyclopedic files correspond to what Perry calls "detached files" (Perry 2001b: 120-1; 2012: 86-9).

[157] "The dominance of non-descriptive information thus entails the essential instability of egocentric concepts. Non-descriptive information about an object is made available to us in virtue of a fundamental epistemic relation holding at time t between us and this object. When, at time t', the relation ceases to hold, this type of information is no longer available. [...] Egocentric concepts are thus temporary object files. They do not last very long. In contrast to egocentric concepts, which are unstable, an encyclopedia entry is a stable object file, because it is not dominated by non-descriptive information. Since the non-descriptive information is not dominant, the object file is made independent of the "context of acquaintance" (i.e. the set of fundamental epistemic relations, with their relata). The existence of the file is no longer relativized to a particular context of acquaintance" (Recanati 1993: 127).

[158] For a critical discussion, see Papineau (2006).

CHAPTER 3

FICTION AND MENTAL FILES

3.1 Real names in fiction

So far, philosophers mainly focused on the ontological and metaphysical analysis of fictional sentences, trying to determine their truth-values. Unfortunately, no theory clearly prevails over the others and the debate seems to reach a deadlock, without concluding in favor of either the realist or the antirealist approach[159]. Perhaps, the conclusion to be drawn is that "if we continue to exploit language to support our own ontological beliefs, sooner or later we end up in a road with no exit" (Voltolini 2010: 126, *my translation*). I partially agree and partially disagree with Voltolini. The mere analysis of sentences can hardly provide a definite answer to this kind of problems. However, at the same time, I also think that, before addressing the discussion on ontology and metaphysics, we should tackle a preliminary cognitive investigation. As Crane said:

> But a theory of what is understood by speakers when they use a name is different from a theory of the semantic value of a name: that is, what contribution a name makes to the truth-conditions of sentences in which it occurs. The Millian theory of names should be thought in this way: it says that the contribution made by a name to the truth or falsehood (the semantic value) of sentences in which it occurs is the fact that it stands for the objects it does. But this is clearly a different kind of theory from a theory about what speakers understand when they use a name. We should distinguish, then, a theory of speakers' understanding of names from a theory of their semantic value. (Crane 2013: 144)

[159] See ch. 1.

In a nutshell, one thing is to wonder about the best metaphysical and ontological account of fictional objects, another thing is to analyze the effects that fictions generate in our mind by adopting a cognitive or, to follow Crane's terminology, a psychological perspective:

> Psychologism [...] looks to the psychological facts for the fundamental explanation of thought. [...] Psychologism can be contrasted with a semantic approach to the psychological, which attempts to give a systematic theory of attributions of thoughts and other attitudes. (Crane 2013: 157)

My aim is to clarify, by adopting the theory of mental files, how we represent in our mind the characters of literature and how we organize information that we draw from our reading or listening to fictional sentences. Only after addressing this preliminary task, we can offer some answers to ontological and metaphysical issues.

I shall begin my investigation on fictional characters by focusing on an intuition that, until now, is not one of the most discussed in this field of study. As Voltolini puts it:

> This is the intuition that fictional works concern not only fictional but also *concrete* individuals, especially actually existing ones. Or, in order for the time being to neutralize any commitment to fictional works, it seems intuitively clear that, over and above existentially creative games, storytelling processes also consist of existentially conservative games in which one makes believe of concrete individuals that they possess certain properties. (Voltolini 2006: 117-8)[160]

Many philosophers think that fictional context suspends reference to real individuals, so that when real names are employed in fictions, they do not refer to anything or they refer to

[160] To put the issue in another way, we can say that, over and above *native characters*, i.e. full-fledged fictional individuals that originate in fiction itself, fictions also involve *immigrant concrete individuals*, i.e. objects that exist in the actual world independently to the fiction. For the notion of immigrant character, see Parsons 1980.

fictional surrogates somehow correlated with their real counterpart[161]. It is not the aim of my survey to discuss the arguments for or against this view. It is sufficient to say that:

> the question of whether there are immigrant concrete objects in fiction concerns only *non-conniving* uses of fictional sentences. Indeed, as far as conniving uses of such sentences are concerned, it is indisputably the case that they may be about concrete individuals. Ordinary existentially conservative games of make-believe typically involve such uses. Since in such games one makes believe of *a certain concrete individual* that it is such and such, one will often make the corresponding linguistic mock-assertion about that very individual. (Voltolini 2006: 118)

When we are engaged into a game of make-believe, the theory that proper names suspend their ordinary reference appears to be out of place, since it collides with our natural way of understanding literary works, when the story is not entirely located in a fantasy world. Let me give an example. When in 1906 Upton Sinclair wrote his novel *The Jungle*, he aimed to denounce, through a fictional story, the exploitation of immigrants in the United States and the harsh living conditions of workers in Chicago, as well as to expose health violations and unsanitary practices in the American meatpacking industry during the early 20th century. His work prompted a public outcry that led to reforms such as the *Meat Inspection Act of 1906*. We cannot explain the influence that this book had on public opinion if we are forced to say that Sinclair's story is not talking about the real Chicago and the real United States, only because he wrote a novel and not a newspaper article. We can now move to consider another book, García Márquez's *One Hundred Years of Solitude*, whose events are located in the fictional town of Macondo. Denunciation of social injustices appears at the core of this book too. We could compare, for example, the living condition of banana workers in the plantations of Macondo with the situation of exploitation in Chicago factories. However, no one would make petitions to change the state of affairs in Macondo. The name "Macondo" is empty, even if there is a town that is very similar to the one featured in *One Hundred Years of Solitude*. We can say that Márquez was inspired by his childhood's town Aracataca, but Macondo does not refer either to Aracataca or to any real place. The description of the banana

[161] See, for instance, Landini (1990), Bonomi (1994, 2008), Lamarque-Olsen (1994), Voltolini (2003, 2006, 2013), Motoarca (2014), Terrone (2017a). Usually, with fictional surrogates, one means fictional entities that intentionally - i.e., owing to the author's choice - correspond to real entities by somehow sharing a significant number of properties with them. For the opposite view, see Friend (2000, 2011).

workers' strike and the massacre that followed is not the report of a specific and historically well-defined case, but a symbol of a situation of injustice and oppression. On the contrary, in order to explain the political effect that Sinclair's book produced, we need to acknowledge that the name "Chicago" in *The jungle* refers to the city of Chicago, even if the name is used into a novel and some fictional predicates are ascribed to the city, such as that it was the place where Jurgis Rudkus and his family, the main characters of the book, used to live.

I agree with Stacie Friend when she claims that:

> "No reference" theories of fiction maintain that fictional contexts suspend reference to real individuals, so that the "London of *1984*" is a fictional place rather than the real location. But this view is false. The fact that in the novel London is portrayed as having properties that it never really possessed should not make a difference, so long as the correct interpretation of the fiction indicates that reference has been made to the real city. (Friend 2011: 192)
>
> The name London plus other cues prompt the imagination, causing those of us who recognize the reference to open our dossier on London and to begin associating our pre-existing London-notions with new fictional information. (Friend 2011: 202)

It is crucial, for the comprehension of *1984*, to recognize that the story is set in London, and to refuse the idea that Orwell simply describes a fictional place homonymous, and in some respects similar, to the real city. This is true for dystopian novels, as well as parodies, historical tales and, in general, whenever a certain real place, person or object is imported into a fiction.

In a nutshell, I assume that – when we engage in games of make-believe and deal with conniving uses of fictional sentences – 1) real names can be employed in fictional stories, and in fact it often happens, and 2) we get fictional information about the referent of that names. When we accept these two claims, some questions immediately arise. How are real names imported in fiction? How do we treat them compared to fictional names? How can we assign fictional properties to real entities? How is fictional information about real entities stored in our mind? I shall argue that the theory of mental files can answer to all these questions.

3.2 Indexed files in non-fictional contexts

How can the theory of mental file help us? The theory, so as presented in chapter 2, seems to be unsuitable for present purposes due to a serious limit: the role of the acquaintance. In fact, mental files framework rests on two principles:

> a) The subject cannot entertain a singular thought about an object *a* without possessing, and exercising, a mental file whose referent is *a.*
>
> b) To possess and exercise a mental file whose referent is *a* the subject must stand in some acquaintance relation to *a*.

Principle B links the existence of mental files to the presence of an acquaintance relation. Together, the two principles entail that no singular thought can be entertained unless the subject is acquainted with what her thought is about.[162] This thesis is also called Strong Acquaintance View (Recanati 2012: 159)[163]. Now, the problem is that we cannot have acquaintance, either direct or mediated, with a literary character. At least, we cannot have the same sort of ER relation we have with a table standing in front of us, or with a friend, or with any other concrete object in the world, because literary characters are not ordinary existing objects.[164] So, the Strong Acquaintance View prevents the theory from being applied in this context. But we are not forced to accept it. Such a strong condition has been criticized by Coliva & Belleri (2013) as too restrictive, since it does not take into account the fact that there

[162] See Recanati (2012: 155).

[163] It is the position defended by Evans (1982: 50).

[164] See ch.1. Fictional antirealists deny the existence of literary characters. On the contrary, fictional realists admit some forms of existence for literary characters, as objects of possible worlds, abstract artifacts or Meinongian objects, but they are not the kind of entities with which an ordinary relation of acquaintance can be established.

are many entities of which, as it seems plausible, we can have a mental representation even in the absence of acquaintance:

> On the face of it, however, this would entail that mental files are quite limited, for we do not seem to be acquainted with a lot of entities that we are nevertheless able to think about. We are not acquainted with non-existing and fictional entities; nor are we acquainted with past or future entities, let alone with abstract ones, like numbers or logical entities.
>
> (Coliva & Belleri 2013: 110)

In this regard, Keith Hall (2013: 119-122) and Peter Pagin (2013: 142) observe that the acquaintance constraint must be understood in a *normative*, non-factual way. Recanati himself confirms this idea:

> I say that mental files are governed by an acquaintance norm. [...] Tokens of a given file carry the presupposition that the norm governing the file is satisfied: they carry an acquaintance presupposition. (Recanati 2013c : 207)

We are thus invited to distinguish between *de facto* and *de jure* conditions for singular thought. Principle B is a *normative* claim, distinct from the *factual* claim that there are no mental files tokened without acquaintance relation to a referent. Given this distinction, we can reject the view that singular thought necessarily requires actual acquaintance and chose a more liberal option according to which as soon as we open a file for an object (whether or not the associated *de jure* requirement is actually satisfied), we put oneself in a position to entertain singular thoughts about that object[165].

But how can we accommodate acquaintanceless singular thought and maintain, at the same time, the idea that mental files are governed by acquaintance constraints, normatively understood? We need to distinguish between *thought-vehicle* and *thought-content*:

> Opening a mental file is sufficient to entertain a singular thought only in the sense of thought-*vehicle*. It is not sufficient to entertain a singular thought in the sense of thought-*content*. What are the conditions on successfully thinking singular thought-*contents*? I

[165] See Recanati (2012: 128). A similar distinction is also stated in *Direct Reference* (Recanati 1993: 178-9) in connection with descriptive names.

> have argued that singular thoughts are fundamentally non-descriptive: their object is determined relationally, not satisfactionally. That entails that one can express a singular thought only in virtue of some relation to the referent. (Recanati 2012: 169)

To think a singular thought in the sense of vehicle, one must activate a mental file. But in order to grasp also a thought-content, reference must be achieved. If reference is not achieved, no singular truth-condition is determined and the thought cannot be evaluated as true or false:

> [...] The content we're talking about is truth-conditional content. A "successful" singular thought is a thought that has singular truth-conditions, that is, a thought such that there is an *x* such as the thought is true (with respect to an arbitrary possible world) iff ... *x* ... The singular content of such a thought is object-dependent: there is no such content if there is no object to which the speaker refers by deploying the relevant mental file.
>
> (Recanati 2013a : 6)

In ordinary cases, thought-vehicle and thought-content go together, since we open mental files in the presence of ER relations that guarantee reference. But there are situations in which the subject can open files even if actual ER relations are not given. For instance, when she expects to be acquainted with the referent in the near future. This is the case of descriptive names like "Neptune", introduced to refer to whatever planet causes certain perturbations in the orbit of Uranus, or "Jack the Ripper", introduced to refer to whoever committed certain murders. Such names are "created only in the expectation that more information about the bearer will accumulate, thus eventuating in the possibility of thinking of the latter non-descriptively. This possibility is simply anticipated by the use of a descriptive name" (Recanati 1993: 180). Expectations may be disappointed, as it happens when the name reveals itself to be empty, like in the case of Vulcan, the hypothetical planed that Le Verrier mistakenly postulated after having observed some peculiarities in Mercury's orbit. By thinking about a non-existing object such as Vulcan, Le Verrier failed to entertain a singular thought in the sense of thought-content, but he was still entertaining a singular thought in the sense of thought-vehicle.[166]

[166] See Recanati (2012: 166) and (2013a: 6). As usual, it is here followed the philosophical tradition in oversimplifying the history of astronomy; see Roseveare (1982).

Expected acquaintance is not the only case in which we can open a mental file in absence of acquaintance[167]. For our purposes, the most interesting examples are provided by *imaginative acquaintance* with fictional objects[168]. Unlike descriptive names, when we think about fictional objects we do not expect any future acquaintance with a referent, for we know that the referent does not exist. Thus, there is no room for mistakes or failures, such as for Vulcan. The norm of acquaintance is explicitly violated, or "exploited":

> All such cases are characterized by the following two features: (i) there is no attempt to satisfy the norm, yet (ii) the norm remains operative. The norm remains operative because it is "exploited", to use Grice's phrase. (According to Grice, the maxims of conversation are exploited when one flouts them blatantly in order to convey an implicature.) What the subject is doing with the file in the relevant cases is not governed by the norm — so the subject cannot be criticized for violating the norm — but doing what he or she does *still presuppose the norm*.
>
> The most general exploitation mechanism is *simulation*: one does *as if* one were using the file normally (i.e. according to the norm) to refer to an object of acquaintance. [...] Does the Emma Bovary file activated in the reader of Madame Bovary refer? The answer ought to be negative: there is simulated reference, but no actual reference. (Recanati 2013c : 2-3)

When we read a novel, for instance Flaubert's *Emma Bovary*, we place ourselves inside a fictional game in which we suspend the rules of everyday life and behave *as if* we had (perceptual or testimonial) acquaintance with Emma Bovary and the other characters featured in the novel. We open a mental file pretending to refer to Emma Bovary, so our thinking is singular in the sense of thought-vehicle, but no singular thought in the sense of thought-content is entertained. There is simulated reference, no actual reference[169].

[167] For instance, we can think about the average mid-twentieth century American, give him a name and predicate things on him. See Breckenridge & Magidor (2012) and Recanati (2013a). See also Jeshion (2010) for the analysis of other acquaintanceless examples.

[168] As Recanati says, "Imagined acquaintance, just as expected acquaintance, justifies opening a file and tokening a singular term in thought" (2012: 168).

[169] "In the Emma Bovary case, the reader makes no mistake and cannot be criticized. But the reason for that is that his or her thought occurs in a special, fictional mode. It does not matter that the files deployed by the reader of a fictional text fail to refer. As Frege suggested, in fiction the step from sense to reference is not taken. In my framework, that means that the primary content is sufficient — one does not care about the secondary content, so it does not matter if there is no secondary content for the fictional thought. Still, according to Frege, what the subject is thinking is neither true nor false." (Recanati 2013c: 208). For a more

The mechanism of simulation offers us a possible account of how we possess mental files to store information about literary characters and nonexistent objects, even though the normal referential condition is not fulfilled. But a further option is available. According to Recanati, the primary function of mental files is to ensure that our mental representation about an object has that particular object as referent. The presupposition of acquaintance is aimed to reach this goal. But, besides their primary function, mental files may also acquire and serve derived functions that do not require acquaintance:

> The singular vehicles, or mental files, we use to cope with simple referential communication (and which have counterparts in the realm of perception), have evolved new, derived functions which enable them to be used also in thinking complex thoughts [...] which, because they involve negation, quantification, conditionals, attitude ascriptions or modals, simultaneously describe situations at several levels.
>
> (Recanati 2012: 176)

In this place, the derived function that I shall take into account is the *meta-representational function*[170]. Mental files have a meta-representational function when they are used to represent the way other subjects think about objects in the world.

Consider the following case: I describe to you a friend of mine, Lorenzo, whom you have never met, as a man of charming personality; a bright, trustworthy, honest guy. In few words, a true gentleman. According to the theory of mental files, you store this information into a vicarious file, a file that is indexed to me, since you are not yet acquainted with the individual I am talking about (you may even doubt that he really exists)[171]. Thus, an *indexed file* is:

---

detailed discussion about the simulation of fictional objects, see Recanati (1998, 2000, 2006b).

[170] For an examination of other derived functions that mental files can play, see Recanati (2012, ch. 13). For an in-depth analysis of the meta-representational function, see Recanati (2012, ch. 14-15) and (2013a). The notion of vicarious files has also been recently applied in cognitive studies to analyze children's developing understanding of identity in alternative naming tasks and belief. See Perner & Leahy (2016).

[171] It should be added that we open an indexed file only when we have reason not to fully trust the interlocutor, or when the situation requires to do so. For instance, a student who first hears about Aristotle at the high school does not open a file indexed to her philosophy professor, but she is justified to create a regular file directly. This is a case of knowledge by mediated acquaintance (see ch. 2.8). I thank Tommaso Piazza for focusing my attention on this point.

> a file that stands, in the subject's mind, for another subject's file about an object. An indexed file consists of a file and an index, where the index refers to the other subject whose own file the indexed file stands for or simulates. (Recanati 2012: 183)

An indexed file, $<f, S_2>$, is thus a file that a subject S_1 uses to represent a files *f* that stands in the mind of another subject S_2 (or in the mind of S_1 in the past time). Unlike regular files, they do not presuppose any norm of acquaintance: indexed files, in fact, do not guarantee to our thought the reference to objects in the real world, since they are mere simulation tools available to "put ourselves in other people's shoes":

> Indexed files, I take it, have an *iconic* dimension. To represent the file deployed by the person to whom a singular attitude is ascribed, we deploy a similar file, indexed to that person. Or perhaps we should say that indexed files are a *simulative* device: by deploying a mental file just like the file in the mind of the indexed person, one simulates the mental state one is attempting to describe; one puts oneself in the another person's shoes (or frame of mind), by looking at things her way. (Recanati 2012: 200)[172]

Then, suppose that you finally meet Lorenzo and you find out that he does not correspond to the person I have described: Lorenzo is a lazy, rude, unpleasant guy. Since you are now acquainted with Lorenzo, a regular file is created in your mind. But you still maintain the previous file, the one indexed to me. It is that file that you deploy in an utterance like:

1) Enrico believes that Lorenzo is a true gentleman.

Attitude ascriptions, and in particular belief reports, are typical cases in which indexed files are used. In fact, "in belief ascriptions, the files associated with linguistic occurrences do not necessarily reflect the speaker's point of view, but may reflect the ascribee's point of view. In other words, attitude ascriptions allow files to be used vicariously." (Recanati 2012: 182). When you attribute to me the belief that Lorenzo is a true gentleman, you are not expressing your own opinion on Lorenzo, but mine. So the file at work is not your regular file, that

[172] See also Recanati (2013a: 6-7) and (2013b: 6).

contains what you really believe about Lorenzo, but a vicarious file expressing my beliefs about Lorenzo.

We can index files not only to other subjects, to represent their own way of thinking about some entities, but also to ourself, to represent how we used to think about an object at a certain time in the past. Recanati discusses the following example[173]:

> 2) We were debating whether to investigate both Hesperus and Phosphorus; but when we got evidence of their true identity, we immediately sent probes there.

According to Recanati, in uttering (2) the speaker currently has only one regular file, since she knows that names "Hesperus" and "Phosphorus" refer to the same planet. But in the past, she used to deploy two distinct files, since she believed that Hesperus and Phosphorus were two separate heavenly bodies. These previous files did not disappear once discovered the identity between Hesperus and Phosphorus, but they remained available as vicarious files indexed to his earlier self, namely a <HESPERUS, self t_1> file and a <PHOSPHORUS, self t_1> file:

> 'Hesperus' and 'Phosphorus' in [2] are associated with separate files whose distinctness reflects the *earlier* doxastic state of the speaker and his group (before they learned the identity of Hesperus and Phosphorus). The point of the sentence is precisely to report the earlier doxastic state of the speaker and his group, and the transition from it to their current doxastic state. (Recanati 2012: 182)[174]

It may be objected that the idea of indexed file is not really necessary to account for attitude ascriptions and belief reports[175]. After all, Recanati's theory appears ontological expensive, since it prescribes, for each object of the external world of which we have a mental representation, a plenitude of files: a regular file that reflects our own point of view, plus an indefinite series of vicarious files that reflect other subject's ways of thinking about that

[173] The original example is from Pinillos (2011: 315).

[174] Vicarious files are also responsible for the opaque reading of belief reports like (2). See Recanati (2012: 185-192).

[175] I am indebted to Stacie Friend for this objection.

particular object. Why should we admit the existence of so many files? An alternative proposal is to conceive only one single file, that contains all information about the object under the form of a list of predicates. Within this list, some predicates are indexed to other subjects. Thus, instead of having a vicarious file <LORENZO, Enrico> containing the information "being a gentleman", you have the indexed predicate <"being a gentleman", Enrico> stored into a regular file LORENZO. In uttering (1), you activate such indexed predicate.

This position may appear attractive, but it is not sufficient to explain some problematic cases, and it turns out that these cases are the most interesting for us. I am talking about those situations in which tokening an empty singular term does not prevent one from expressing a truth-evaluable content[176]. A first example is given by negative existentials that seem to express true propositions, even if they contain empty names. An example is:

3) Vulcan does not exist

Another case is provided by belief reports that ascribe to another subject what Recanati calls a pseudo-singular belief, i.e. a belief that has the form of a singular mental sentence in the sense of thought-vehicle, but not in the sense of truth content, since it does not express any proposition[177]:

4) Le Verrier thought that the discovery of Vulcan would make him famous

5) My son believes that Santa Claus will come tonight

The name "Vulcan" is, at least for the speaker, an empty names, so she has not a regular file to deploy in uttering (4). Since there is no regular file, there cannot be an indexed predicate <"making Le Verrier famous", Le Verrier> stored in it. The same for sentence (5): there is no regular SANTA CLAUS file in which to insert the predicate "being coming tonight" indexed

[176] I thank François Recanati for emphasizing this point. See Recanati (2013a : 6).

[177] "To entertain a pseudo-singular belief is to have a singular mental sentence tokened in one's belief box, but one that fails to express any proposition." (Recanati 2012: 177, *footnote*). See also Recanati (1998: 557; 2000: 226).

to my son. In this situation, we are not even legitimated to invoke a mechanism of simulation, since it is assumed that (4) and (5) are uttered in a serious, non-conniving, way, with the intention to tell genuine truths from a perspective that transcends the context of make-believe. The speaker is not merely pretending to report her son's expectations about Santa Claus, she is truly reporting what the child believes. The idea that we are not playing a game of make-believe is reinforced once we consider negative existentials. By uttering (3), we are not simulating some kind of acquaintance to a nonexistent planed called "Vulcan", but we are just denying that the name "Vulcan" refers to anything: "For when we say that Vulcan doesn't exist, we surely take ourselves to be denying that there is such a thing rather than affirming that there is such a thing but that it is unreal" (Sainsbury & Tye, 2012: 16 – 17).

Finally, difficulties become more serious when we consider the phenomenon of intentional identity: "we have intentional identity when a number of people, or one person on different occasions, have attitudes with a common focus, whether or not there actually is something at that focus" (Geach 1967: 627). In Geach's famous example, a reporter, who is not superstitious, describes an outbreak of witch mania that affects the inhabitants of a country village. The journalist writes the following sentence:

> 6) Hob thinks a witch has blighted Bob's mare, and Nob wonders whether she killed Cob's sow

Although the reporter does not believe in witches, therefore not entertaining any regular file for them, he can successfully ascribe to the inhabitants thoughts about a certain witch who did actions like blighting Bob's mare and killing Cob's sow. Under which files should we store these pieces of information? Unlike standard cases of belief reports, like (4) and (5), we have here an additional problem. Not only we ascribe beliefs by using an empty singular term, but we presuppose that this singular term is shared by different people and all of them have thoughts concerning the same focus.

Recanati's theory does not suffer from these limits because it is possible to open vicarious files even in the absence of the regular one. There are two possibilities for a given indexed file:

> Either the indexed file, which represents some other way of thinking about some entity, is linked to some regular file in the subject's mind referring to the same entity (and corresponding to the subject's own way of thinking of that entity); or it isn't. If it isn't, the subject only access to entity in question is via the filing system of other subjects.
>
> (Recanati 2013a : 8)[178]

In the first case, indexed files are *loaded*, i.e. linked to regular files. But when it happens, the type of connection is different from the one that is established when two regular files are linked[179]. Linking between regular files allows for a free share of information, whereas as long as single files remain separate, one cannot integrate the data stored into them. Normally, the subject creates the link due to a discovery of identity between the reference of the files: she comes to know that Hesperus and Phosphorus are the same planet, so all information that she learned about Hesperus can legitimately be attributed to Phosphorus and vice versa. But with vicarious files the situation is different:

> There is an important difference between linking as it operates between regular files (horizontal linking), and linking as it operates between regular files and indexed files, or between indexed files of different degrees of embedding (vertical linking). *Linking between regular files typically makes it possible for information to flow freely between the linked files.* But indexed files are used to stand for some other subject's body of information about some object, and that function could not be served if, through linking, the indexed file was contaminated by the subject's own information about that object. [...] *So vertical linking between regular files and indexed files (or between indexed files with different degrees of embedding) preserves the informational encapsulation of files*, which standard (horizontal) linking typically has the effect of suppressing.
>
> (Recanati 2014: 475-476, *my italic*)[180]

As we said, *horizontal linking* is the standard type of linking for regular files. Conversely, *vertical linking* typically concerns indexed files and has the function to preserve the data encapsulated in each file. In a belief report like (1), it is the ascribee that believes that Lorenzo

[178] See also Recanati (2012: 184).

[179] See ch. 2.9.

[180] See also Recanati (2012: 184, 193).

is a true gentleman, not the speaker: the predicate "being a true gentleman" belongs to the vicarious file, not to the speaker's regular one[181]. However, even if the indexed file does not join the data with the regular one, it acquires, thanks to this connection, the referent that the regular file has. By uttering (1), the speaker refers to Lorenzo, besides the fact that she is reporting someone else's beliefs about him.

Given the existence of two types of files in the subject's mind and the mechanism of linking that operates between loaded vicarious files, the subject:

> has two ways of thinking of the object: a way of thinking of his own (a regular file) and a vicarious way of thinking (the indexed file). If the subject uses the indexed file to think about the object, that use is "loaded" and has existential import. [...] Even though the subject refers to the object through some other subject's file about it, he takes that object to exist since he himself has a regular file about it. (Recanati 2012: 184-185)

This is quite useful in everyday life. In fact, people's representations about things of the world often differ very much from each other and by means of loaded index files we can represent all these different points of view. We can take the divergence of opinion about politicians as a paradigmatic case[182]. Suppose that two journalists, Emilio Fede and Marco Travaglio, write a biography about an Italian political man, Silvio Berlusconi. It is well known that Fede admires Berlusconi whereas Travaglio is one of his fiercest critics. Not surprisingly, by reading their books, we obtain two very different portraits of the same individual. Since Berlusconi is a real person and we have direct or mediated acquaintance with him, it is natural to compare these dissimilar representations with our personal opinion about Berlusconi. In other words, we compare information contained in files <BERLUSCONI, Fede> and <BERLUSCONI, Travaglio> with information stored in our primary file BERLUSCONI. As a result of this operation, we may conclude that Fede's book is too much celebratory, while Travaglio's one is excessively derogatory. I must add that it makes sense to wonder whether Fede's or Travaglio's biographies are faithful to reality not only because the name "Berlusconi" has reference, but also because we take the books as nonfiction. We are

[181] Vertical linking also allows importation and exportation of notions within and outside the fiction. I shall return on this point later (ch. 3.3).

[182] I thank Carola Barbero for the discussion about these themes and the example she suggested to me.

supposed to evaluate their statements on Berlusconi as true or false. We act in the same way when we read biographies about historical figures of the past, like Julius Caesar or Napoleon.

Let us now consider the case in which indexed files are not linked to any regular file. We have what Recanati calls a *free-wheeling*, or *unloaded*, use of indexed file (2012: 184). In this situation, the subject can only think about an object via the filing system of other subjects: all information at her disposal are those stored in the vicarious file. Contrary to what is affirmed by some theorists (Perner & Brandl 2005; Perner, Rendl & Garnham 2007), Recanati holds that "an indexed file cannot be directly anchored to a real object, but only via a regular file to which the indexed file is linked." (Recanati 2012: 184, *footnote*). Unloaded files do not refer to anything; they simulate a referent, thereby exploiting the phenomenological effect of the acquaintance constraint. Therefore, the more natural field of application for this kind of files turns out to be empty names. Two options are available. In the first case, we are using empty names within a context of pretense. For example, if I utter:

7) Santa Claus will come tonight

with a conniving use, sentence (7) has merely fictional truth-values. I am only pretending to refer to Santa Claus, and pretending to predicate something of him. By uttering (7), I am not expressing any genuine singular thought but, to use Frege's words, a *mock thought*[183]. Here, "the file associated with Santa Claus is indexed to Santa Claus's believers and unloaded, so the whole speech act has to be seen as a form of pretense." (Recanati 2012: 204).

The second option involves cases such as the ones we have seen before: negative existentials like (3) and beliefs reports like (4) and (5). According to Recanati:

> the utterance expresses a thought that is globally meta-representational – it is about someone's, e.g. my children's, representations, rather than about what these representations are about. This corresponds to pseudo-belief ascriptions ("My children believes that Santa Claus will come tonight"). Negative (and positive) existentials arguably work the same way: the file associated with the singular term in "Vulcan doesn't exist" is a free-wheeling indexed file, and the statement is globally meta-representational (it says that the file does not refer). (Recanati 2012: 204)

[183] Frege (1897: 130).

Other interesting cases of application for unloaded vicarious files are intentional uses of empty names or singular terms that involve Geach's phenomenon of "intentional identity", like in sentence (6):

> 6) Hob thinks a witch has blighted Bob's mare, and Nob wonders whether she killed Cob's sow

Following Recanati's analysis, (6) is a special case of pseudo-singular belief ascription since, at first, the speaker ascribes to Hob the belief that a witch has blighted Bob's mare, and then she ascribes to Nob the suspect that the same witch killed Cob's sow. The peculiarity of this example consists in the fact that "the anaphoric relation between the indefinite and the pronoun implies that *the same file* is deployed by Nob and by Hob, even though no communication takes place between Hob and Nob." (Recanati 2012: 205). For an exhaustive account of intentional identity, vicarious files are needed, but not sufficient. We should introduce the notion of public files, i.e. files shared by distinct individuals in a community.[184] Nowadays, there is much debate about public files and, more generally, about intersubjective communication involving empty names. In literature, two main strategies are proposed: according to Friend (2011, 2014), that is inspired by Evans (1973) and Perry (2001a), the dominant source of information about a certain empty name determines practices that underpin our uses of that names (*info-centric approach*), while according to Sainsbury (2005), Salis (2013), and García-Carpintero (2016a), that are inspired by Kripke (1980), people's practices retain the original intention of who firstly introduced the name with an act of baptism (*name-centric approach*). For the purpose of this discussion, it is sufficient to observe that both positions are compatible with Recanati's theory.[185]

[184] For the notion of public file, see Perry (2002). For a recent discussion on this topic, see Terrone (2017b).

[185] See Lo Guercio (2015).

3.3 Indexed files in fictional contexts

According to Recanati's theory, we can think about nonexisting objects by using two different types of mental files, in both cases without reference to any entity. In the first case, we have regular files that presuppose the acquaintance constraint, but explicitly violate this condition through a mechanism of simulation. So, when we read Flaubert's novel *Madame Bovary*, we engage in a fictional game in which we suspend the rules of everyday life and we act as if we are acquainted with the characters described by the book. In the second case, we use indexed files with meta-representational function[186]. This is what happens when we consider a sentence like:

4) Le Verrier thought that the discovery of Vulcan would make him famous

Sentence (4) ascribes, with success, a pseudo-singular belief to Le Verrier[187]: we attribute to Le Verrier a thought that has a singular form, but that does not express any proposition at all[188]. We are thinking about the representation of someone else and not about the reference of that representation.

As regards mental files applied to nonexisting and fictional objects, Recanati does not push his analysis further, leaving many questions unanswered[189]. The theory depicted up to now is a good starting point, but it is not sufficient. My aim is therefore to develop a more complex account taking inspiration from the base that has been laid. I take for granted the

[186] I leave aside the case of Fregean mock thoughts.

[187] "He [the ascriber] can exploit the Meinongian pretense and do as if there was such an object as Martins, Holmes or Santa Claus; that is, he can fictively ascribe to the believer a singular belief concerning Martins, Holmes or Santa Claus. Since it is clear that the pretense is pretense and that, in facts, there is no such individual, the fictive ascription of a singular belief concerning that individual amounts to the factive ascription of a pseudo-singular belief – an ascription which [...] it is not directly expressible save by appealing to the pretense." (Recanati 2000: 226). See also Recanati (2012: 204).

[188] Or, at most, a *gappy proposition*. See Everett (2003).

[189] However, for a recent development of the theory in this direction, see Recanati (2018). For some critiques to Recanati's notion of pseudo-singular thought, see Crane (2013: 158-162), Ninan (2015), Stojanovic & Fernandez (2015).

distinction between regular and indexed files, but I think that the latter are much more useful for understanding fictional objects. Therefore, I shall concentrate mainly on this second notion.

So far, we have seen that indexed files may be loaded or unloaded. Unloaded files are typically associated with empty names, whereas loaded files are linked with regular files, inherit their referent and allow us to figure out how other subjects think about objects of the word. So it seems that loaded files have no place in fictional contexts. But, I claim, this is wrong. As we said at the beginning of the chapter, in fictional stories it is quite common to find names of real people, places, and other things. Thus, the question I wish to address is: what happens when we deal with real names within a fiction? I suggest that the situation is not dissimilar to when we read a biography or a history book about some famous flesh-and-blood man. Let us recall the example of the two journalists, Emilio Fede and Marco Travaglio, that write a book on Berlusconi, giving different portraits of the same individual. We generate two separate indexed files. Since they are linked with our regular one, we may compare our own knowledge about Berlusconi with information drawn from the books. When we read a novel displaying the name of a real person, we acquire new information about that individual. It does not matter that information is fictitious: we produce a vicarious file and we keep information encapsulated in it. Since the vicarious file is vertically linked with our regular file, their content is not joined: we do not mix fictional information of the former file with real information stored in the latter. Both files can be exploited, depending on our purposes. We can import in the fiction our knowledge about people or places the story is about, if the story allows or invites us to do so. For example, once we recognize that Orwell's *1984* takes place in a dystopian London in the near future, we open our regular file on London and associate "our pre-existing London-notions with new fictional information", as Friend suggests (2011: 192). The idea of importation is supported by those philosophers who adopt the *Reality Principle*, or similar ones, to investigate what is true in a fiction:

> (RP) If p_1, ..., p_n are the proposition whose fictionality a representation generates directly, another proposition, q, is fictional in it if, and only if, were it the case that p_1, ..., p_n, it would be the case that q.[190]

[190] Walton (1990: 145). Other classic formulations are in Lewis (1978), Ryan (1980), Wolterstorff (1980) and Evans (1982).

As Kendall Walton says:

> The basic strategy which the Reality Principles attempts to codify is that of making fictional worlds as much like the real one as the core of primary fictional truths permits. It is because people in the real world have blood in their veins, births, and backsides that fictional characters are presumed to possess these attributes. [...] The interpreter is to ask what the real world would be like if the proposition whose fictionality is generated directly were true: What else would be true if they were? The answer gives the propositions whose fictionality the primary fictional truths imply. (Walton 1990: 144-145)[191]

Conversely, we can export notions that we gain from novels in our regular files, when we recognize such body of information as non-fictional. A historical novel may depict with high accuracy not only places, events and notable historical figures, but also ways of living, manners, social conditions and other details of the past. For instance, Manzoni's *I Promessi Sposi* is based on rigorous researches and by reading it we learn something true about the 17th century, such as the story of the nun of Monza (Marianna de Leyva y Marino) and the events of the Great Plague of 1629-1631, reconstructed on archival documents and chronicles of the time. Or, by reading Sinclair's *The Jungle*, we come to know about the exploitation of immigrants in the United States as well as the harsh living conditions of workers in Chicago farms at the beginning of 20th century.

Finally, we can make a comparison between fiction and reality, for example by turning around the city of London at the research of places mentioned in Conan Doyle's stories. We may have fun visiting the Sherlock Holmes Museum at 221B Baker Street, even if we know that at that time the street numbers only arrived until 85. As an alternative, we may examine the reliability of historical figures depicted in fictional novels. But here a clarification it is needed. One may wonder whether Shakespeare provides a faithful account of Caesar in his drama, or whether the figure of Napoleon described by Tolstoy's *War and Peace* matches the reality. However, these considerations are subordinate to other types of evaluation, such as those concerning aesthetics and literature. Since we are engaged in a game of make-believe,

[191] Friend has recently argued in favor of a *Reality Assumption*: "everything that is true or obtains in the real world is storified—that is, we are invited to imagine it as part of the storyworld—unless it is excluded by the work" (2017: 5). By contrast with the more familiar Reality Principle, the Reality Assumption has the merit of being more general and answering for some critics raised against the Reality Principle.

we are ready to accept information about Caesar or Napoleon without inquiring too much about its truth. We do not blame Shakespeare for inventing some events, because we expect him to do so, whereas our reaction would be very different if we discovered that Fede or Travaglio invented something about Berlusconi. Shakespeare's drama is not considered a reliable source of information, contrary to Fede's or Travaglio's books, but a source of fictional notions about Caesar.

So far, philosophers have discussed a lot whether we can obtain true knowledge by reading fictional works and, when they give a positive answer, which kind of knowledge it is[192]. I wish to underline that, if we accept the possibility to import notions from the real world within the fiction and to export at least some form of knowledge from the fiction into the real world, the idea of vicarious files proves to be very useful. In fact, vertical linking between vicarious and regular files easily explains in which way we keep separate real and fictional information, even in case of non-empty names used in fictional stories, and how knowledge moves from reality to fiction and *vice versa*.

What happens, instead, with empty names used in fiction? Starting from Frege[193], philosophers looked at empty names as something exceptional that asks for exceptional explanations, i.e., explanations that differ from the standard account of names having reference. Consequently, they set up the study of fiction as a separate domain of language in which special rules are at work. This also applies to Recanati. He claims that, when we are not attributing pseudo-beliefs to someone else or uttering negative existentials, we deploy regular files. But these files explicitly violate ordinary rules, since they are generated by means of a mechanism of simulation that overcomes the acquaintance constraint. I do not agree with this idea. My suggestion is to turn the perspective upside down. The analysis carried out so far was aimed to show that fictional discourse has nothing special: it can be explained with normal tools made available by the theory of mental file. We are not forced to admit exceptions, since we already have everything we need. Vicarious files find their application regardless of the context of pretense[194], but they can be employed successfully even to explain the way of functioning of real names, like “Caesar”, “Napoleon” or “London”, in fiction.

---

[192] For an introduction on the debate about what we can learn from fiction, see Carpintero (2016b) and Ichino & Currie (2017).

[193] See Frege (1892) and (1897).

[194] See ch. 3.2.

So, what happens when we deal with names of purely fictional character? My claim is that, once again, files work in the same way: we open an indexed file and we begin to store information in it. But in this case, it is a free-wheeling, or unloaded, use of the indexed file, since we lack a regular, acquaintance-dependent file in our mind to be linked to. Thus, no reference to objects of the real word is made and no comparison between fiction and reality is possible. We only have the (fictional) information that we gain from the story and that we store in the indexed file. To grasp this point, just consider the case in which the same fictional character appears in two distinct literary works, being portrayed in different manners depending on the stories. A famous example is Ulysses so as presented in *Odyssey* and in *Divina Commedia*. Since it does not exist any real person that is the referent of the name "Ulysses", and thus of the mental files, it does not make sense to ask whether or not Ulysses in Homer's poems resembles a historical Ulysses more than Dante's one. Information corresponding to the predicates that each work of fiction assigns to Ulysses is gathered in two distinct indexed files and remains encapsulated within them. We can compare these dissimilar versions of Ulysses with each other, assign our preference to one rather than the other, but that is all.

Given this background, it seems that no mechanism of importation or exportation can take place, such as those described for Napoleon or Julius Caesar. I agree as regards the phenomenon of exportation, because there is no further mental box to which information can migrate. However, we can make room for importation, if we accept the Reality Principle, or some variants of it as the Reality Assumption proposed by Friend (2017). According to RP, since the myth tells us that Ulysses is a man, it is reasonable to believe that, in the fiction, he has all the features that human beings normally possess, even if such individual does not really exist. As we have said: "It is because people in the real world have blood in their veins, births, and backsides that fictional characters are presumed to possess these attributes" (Walton 1990: 145). These pieces of information do not come from a regular file ULYSSES, that does not exist, but from our knowledge on human beings, plus other general notions about the way the world is made[195].

[195] How can we account for general notions like "human being"? Probably, we need to make room for the idea that, in addition to singular mental files, i.e., files that refer to particular objects of the external world, there are also general mental files for groups of objects and abstract concepts. The thesis is supported by Crane (2013) and Sainsbury & Tye (2013).

3.4 To what mental files are indexed

I said that we store fictional information about real or fictional characters in indexed files. It is necessary to specify to what mental files are indexed. One can browse several candidates. According to Recanati, vicarious files are used to represent other subject's way of thinking about some object. So we need to individuate a subject, whose way of thinking allows us to access the character in question. The most obvious option is to index vicarious files to the author of the story. For instance, we owe to Conan Doyle for all we know about Sherlock Holmes: his physical aspects, his way of thinking and investigating, his actions, and so on. Sherlock Holmes is a figment of his author's imagination, and it is the author who engages with us, the readers, a game of make-believe in which the brilliant detective solves difficult crime puzzles. We could go even further and make an analogy between Doyle's novels on Sherlock Holmes and Travaglio's or Fede's non-fictional books on Berlusconi. In both cases, we store information within indexed files. The difference is that Berlusconi is a real existing person, so we also have a regular file and compare the picture presented by Travaglio or Fede with our mental representation of that man, whereas Sherlock Holmes is just a fictional character, so no further information can be added apart from what is said by Conan Doyle. Moreover, such a perspective looks attractive because it manages to provide a straightforward explanation of why different versions of the same fictional character may exist. As two journalists can give different portraits of the same person, two novelists can produce dissimilar representations of the same character, such as Ulysses described by either Homer or Dante. But since there is no real Ulysses, each author is free to depict his hero as he prefers. No constraint of acquaintance, no rule of faithfulness is at state.

Unfortunately, at a more careful analysis, indexing to the author turns out not to be a very good idea. In fact, it does not allow us to distinguish between fictional and non-fictional sentences or, to be more precise, between sentences on a character that the author utters with a conniving use, i.e. within the context of pretense, and sentences that the same author can utter outside the game of make-believe. What the author really thinks about a particular character

may differ from the content of his books. For instance, it is known that Agatha Christie hated Hercule Poirot and she thought that he was a detestable character, but Poirot does not appear detestable inside the fiction[196]. Even more radically, we can say that Agatha Christie, and any other healthy mind person, never seriously believed Poirot to be a flesh-and-blood man. But, of course, Poirot is a flesh-and-blood man in Christie's novels. I take this argument to be decisive and, thus, I exclude the idea to index mental files to the author of the story.

Another fascinating option is to say that, since we are playing a game of make-believe in which we assume that someone tells us a story as if all events were true, mental files must not be indexed to any real individual, such as the author, but to the fictional storyteller[197]. The narrator may be an anonymous, omniscient and stand-alone voice that merely reports the events, like a journalist. Or it may be a character directly involved in the story, a witness that tells us the facts from his particular point of view, as Watson does in Doyle's novels. In this second case, we would have an indexed file of the type <SHERLOCK HOLMES, Watson>, that involves the character Sherlock Holmes so as presented by his fictional friend Dr. Watson. Similarly, we would have a file <WATSON, Watson> for the character Dr. Watson so as he is described by himself. This solution has several advantages. First of all, it easily overcomes the problem mentioned above. We may have two distinct files: a file <POIROT, Christie> for the mental representation of detective Poirot that Agatha Christies seriously used to have in her real life (containing information that Hercule Poirot is a detestable character, that he is a fictional character, and so on) – and a file <POIROT, Christie*> for the detective Poirot so as featured in the fiction. As a second point, we account for the different versions of the same character that moves from one work of fiction to another, changing its features. There are as many Ulysses as there are storytellers who tell stories about him. Again, it does not make sense to question which Ulysses is the real one, given that Ulysses does not exist. The third, and more interesting, advantage is that we can apply the same analysis to handle with inconsistencies that sometimes affect fictions. A famous example is provided by Watson's wound that "migrates" from the leg (when you read *A Study in Scarlet*) to the shoulder (when you read *The Sign of the Four*). According to the theory of mental files, more than one file can be indexed to the same person, and this also happens in non-fictional contexts. I recall a case already discussed in this chapter[198]:

[196] I am indebted to Diego Marconi for emphasizing that point.

[197] For a discussion on the role of the storyteller for the comprehension of fictional stories, see Currie (2010).

[198] See 3.2.

2) We were debating whether to investigate both Hesperus and Phosphorus; but when we got evidence of their true identity, we immediately sent probes there.

According to Recanati's analysis, in uttering (2) the speaker “knows that Hesperus is Phosphorus; he is aware that there is a single planet, Venus. So he should associate the same mental file [...] with both 'Hesperus' and 'Phosphorus'.” However, “ 'Hesperus' and 'Phosphorus' in [2] are associated with separate files whose distinctness reflects the *earlier* doxastic state of the speaker and his group (before they learnt the identity of Hesperus and Phosphorus)” (Recanati 2012: 181-182). Here, the speaker uses two vicarious files indexed to her earlier self, namely a <HESPERUS, self t_1> file and a <PHOSPHORUS, self t_1> file, both of which are linked to her current and regular file VENUS.

We can adopt a similar analysis for Watson's wound. We store information gained by reading *A Study in Scarlet* in files indexed to the storyteller, i.e. Dr. Watson, at a time t_1. In the sequel *The Sign of the Four*, the narrator still is Dr. Watson, but at time t_2, so we obtain files indexed to the same storyteller at the time t_2. Inconsistency regarding the location of Watson's wound is thus solved since we have in mind two distinct files <WATSON, Watson t_1> and <WATSON, Watson t_2>. The first one is the representation that Watson gives of himself at time t_1, when he was fictionally writing *The study in scarlet* and declared that he was wounded in the shoulder at the Battle of Maiwand by a jezail bullet, the second one is the representation that Watson gives of himself at time t_2, during the drafting of *The Sign of the Four*, when he complains that his leg, hurt by a bullet some time before, aches at every change of time. Since Dr. Watson is not a real individual, it is not necessary for these two files to present a coherent portrait of the character. In other words, there is not a primary file WATSON to which the indexed files are linked. As we cannot compare information stored in the two indexed files with our knowledge about a flesh-and-blood man, both the two versions of Dr. Watson are equally legitimate. For contradictions to disappear, it is sufficient to split the official stories of Sherlock Holmes into two distinct fictions and store information about Watson's wound into the corresponding separate files.

Even this stance has to deal with serious objections. One is that the storyteller may change during the course of the narration. Some stories may have multiple narrators to illustrate the storylines of various characters at the same or different times, or to allow a more

complex, non-singular point of view, by presenting multiple different perspectives about the events. An interesting, and extreme, case of multiple narrators is provided by Orhan Pamuk's novel *My name is red*, in which each chapter is narrated by a difference fictional character. One can answer to this objection by claiming that, although there is a plurality of narrators, we can always appeal to an hypothetical implicit narrator who reports the tales of the other characters. I see nothing fundamentally wrong with this reply, but it does not appear to be a very persuasive explanation and can be considered an *ad hoc* argument. Another problem is related with the case of unreliable narrator[199]. When we have a storyteller whose credibility, for several reasons, has been seriously compromised, we are forced to put in doubt his words and actively interpret the story. But then, not all information stored in mental files can be indexed to the storyteller: on the one hand, we do not accept some bits of information since we do not consider them trustworthy; on the other hand, we add new information that derives from our understanding of the story but it is not directly mentioned in the narration. Finally, there is an even more serious objection. Although there are good reasons to think that the presence of a storyteller is a required element for novels and, in general, for most literary works, either fictional or not[200], in other types of narratives, such as theatrical dramas, television series, video games and movies, it seems more plausible to say that the storyteller is merely optional[201]. Of course, one can restrict the analysis only to traditional literature, excluding those genres of narrative in which the presence of a storyteller is controversial. But this move seems arbitrarily reductive. Indexed mental files should be able to explain how our mental representation of fictional characters works, independently from the way and the medium the stories are implemented with.

One last option remains. We do not need to index mental files to any individual subject, either real such as the author or fictional such as the storyteller, since we can index them directly to the fiction itself. It may be objected that, according to the theory of mental files, indexed files are tools that we have at disposal for representing, in our mind, the point of view of other people. An indexed file simulates the mental state of the indexed subject, so it does not make sense to index files to the fiction, for fictions are not that sort of things that have a mental life. However, I see no theoretical obstacle in stretching the notion of indexed

[199] For a discussion on unreliable narratives, see Currie (1995).

[200] See Currie (2010: 65-85).

[201] I thank Enrico Terrone for emphasizing this objection during a presentation of a draft of my work at the Institut Jean Nicod. For a discussion on this topic, see Wartenberg (2015) and Thomson-Jones (2016).

file so to include also such kind of cases. When we take part in a game of make-believe, we are urged to imagine a certain situation and to adopt certain mental attitudes, for instance, by accepting the told story as unquestionable true[202], no matter whether it involves nonexistent people and events that are bizarre and unrealistic. In the files we store information that we associate to the world of the story, as participants of that game. If we wish to be more precise, we could say that, by indexing a file to a fiction, we mean to participate to a certain practice of make-believe, in which we put ourselves in the mind of someone that is not pretending, but actually believes in the facts depicted by the story[203]. Similarly, in normal, non fictional, situation, we put ourself in the mind of an external subject to represent her mental states. It is just the same act of simulation.

This solution does not suffer from the objections we have seen before. Even if many storytellers are at work, like in Pamuk's *My name is red*, each of them is telling us segments of a single story, so it makes sense to index information in just one file. Similarly, the narrator may be untrustworthy, but since in the mental file we collect solely those pieces of information that we accept as true in the story, we do not passively store all that is told to us. In types of narratives in which the role of the storyteller is not required, like movies or theatrical dramas, we gain information about the story directly by enjoying the fiction, without the need of a filter. Still, one thing is the way a character is depicted inside a story, like Poirot being a flesh-and-blood detective, another thing is the character so as conceived by her author outside the fiction, like Poirot being a detestable fictional character for Agatha Christie. But, maybe, the most interesting advantage of this perspective is that it can throw light on some important puzzles that concern the identity of fictional characters. Next paragraphs will be devoted to such investigation.

[202] With the exception of unreliable narratives.

[203] Alternatively, we could index files to participants of a game of make-believe. I thank Recanati for this suggestion.

3.5 Puzzles on the identity of fictional characters

It is not unusual, especially when a certain literary character is widely known, to find many different versions of it: the same character can move from one work of fiction to another, by appearing in very different stories (sequels, remakes or parodies), and even migrate from one media to another (from a book to a comic or a movie, from a text to a picture to a sculpture)[204]. I quote what follows from Voltolini:

> Pinocchio is not only the protagonist of Collodi's novel, but we also find it in Giovanni Giraldi's book, *Il figlio di Pinocchio*, and even in cartoons, songs and movies (from Walt Disney animations through Roberto Benigni's movie up to the irreverent *Le avventure erotiche di Pinocchio* by Corey Allen). The intuition here is that there is one and the same character that persists, through its changes in the transition from one work to another. (Voltolini 2010: 77, *my translation*)

Migration of fictional characters elicits *prima facie* conflicting intuitions. On the one hand, a fictional character seems linked to the particular work of fiction (a novel, a poem, a movie, etc.) in which it appears: Pinocchio is depicted in one way in Collodi's novel, in another way in Walt Disney's cartoons, and in a still different way in Benigni's or Corey Allen's movies. It is natural to distinguish Collodi's Pinocchio from Disney's or Benigni's ones, since each of them has specific properties. On the other hand, we have the strong temptation to think that Pinocchio is a single fictional character, despite its changing of features in the passage from one work to another. But cases of fission and fusion among fictional characters appear to challenge this intuition. Let us see in detail the problem in its various ramifications.

In order to deal with the identity of literary characters, it is important to recall the distinction between conniving and non-conniving uses of fictional sentences. According to Evans (1982: 365-6), the utterance of a sentence is conniving when the utterer is engaged in a practice of make-believe and the truth-values of the sentence are merely fictional. A *non-*

[204]See ch. 1.4.3.

conniving use is when the sentence is uttered with the intention to tell genuine truths that transcend the context of pretense. As Voltolini observes, non-conniving "uses are intended to enable people to speak *about* the fiction rather than *within* the fiction" (2006: 118). So, a single sentence can be employed with different uses, according to the situation and the speaker's intention. Let us see why this distinction is useful. We can question the identity of a fictional character from a conniving perspective. Sometimes the task is less obvious than it may seem, as the story can be deliberately misleading. A clear example is Stevenson's *Strange Case of Dr. Jekyll and Mr. Hyde*. In the book, Dr. Jekyll and Mr. Hyde are represented in very different ways as two distinct characters and, nonetheless, at the end of the story the author reveals that they are the same person. Thus sentence:

8) Jekyll is Mr. Hyde

states an identity that is stipulated by Stevenson within the fiction.

We can also wonder about the identity of a literary character from a non-conniving perspective. It usually happens when the character appears in different works of fiction. Pinocchio is an example. Ulysses is another good one. Many poets and novelists were inspired by this character. In Greek epic poems, under the name of "Odysseus", he is one of the heroes fighting in the Trojan War and the unlucky traveler who tries to come back to his homeland Ithaca. Even Virgil mentions Ulysses in *Aeneid*. After centuries, he appears as a damned soul that Dante meets on his journey to *Hell*. Even in recent time, his fame does not decrease. James Joyce suggests us to see in Leopold Bloom a new, modern, Ulysses. Given that it presupposes the comparison between (at least) two separate fictions, the following identity statements must be considered as a metafictional sentence, i.e., a sentence uttered with a non-conniving use:

9) Homer's Odysseus and Dante's Ulysses are the same fictional character

Evaluating this sentence is much less easy than (8). As already said, on one side we have the intuition to deal with the same character, although the various stories in which it appears assign to it different properties, sometimes in contrast to each other or even contradictory. On

the other side, we have another equally strong intuition that, in a certain way, there is an Ulysses belonging to the *Iliad*, one belonging to the *Odyssey*, another to the *Aeneid*, yet another to Dante's *Divina Commedia*, and so on. The same goes for many other characters of the literature.

Now let us focus on a more complex case: the famous dispute between Cervantes and Avellaneda. The first part of Cervantes's *Don Quixote* was published in 1605. The success of the novel was such that an anonymous author, under the pseudonym of Alonso Fernández de Avellaneda, released a sequel in 1614. Cervantes then decided to write the continuation of Don Quixote's story, published in 1615. In the *Preface* that introduces this second part of *Don Quixote,* Cervantes criticizes the spurious sequel and promises to conclude all hidalgo adventures until his death and burial, therefore discrediting Avellaneda's work. Thus, how should the following sentence be considered?

10) Avellaneda's Don Quixote and Cervantes's Don Quixote are not the same character

Compared to the previous examples, is (10) a fictional or a metafictional sentence? It presupposes the comparison between two different texts on Don Quixote, but at the same time is a sentence that claims something internal to Cervantes's story. As I shall explain later, this is a case in which the same sentences can be uttered either with a conniving or with a non-conniving use, and the type of analysis that (10) receives changes according to the use in question.

Lastly, an account of the identity of fictional characters should be able to treat problematic cases such as phenomena of fusion and fission. This point is a serious obstacle for some theoretical positions, such as the artifactualist account proposed by Thomasson (1999)[205]. In this regard, Voltolini observes:

> [...] the artifactualist maintains that treating fictional characters as abstract artifacts legitimates the intuition that *one and the same character* migrates from one work to others. But cases of fusion and fission show that this intuition is not as well founded as it seems. If those who were two characters become one and the one who was a single character doubles, how can we guarantee that, even when such cases do not occur, there is a

[205] See ch. 1.4.3.

> relationship of numerical identity between the character of a story and the (relatively) similar character of another story? (Voltolini 2010: 84, *my translation*)

A case of *fusion* is described by Bonomi (1994: 66). In an early version of *À la recherche du temps perdu*, Proust presents two distinct characters, the musician Berget and the naturalist Vington, whereas in the ultimate version of the *Recherche* there is only the musician Vinteuil, who somehow sums up the features of the two. We would say that Proust imports both Vington and Berget into the later version, by fusing them into a single character. *Fission* is just the reverse process. In Carroll's *Alice in Wonderland*, the character of the Duchess is an antagonist of the Queen of Hearts, but both originate from the split of a preexisting character. In the original draft of the story, *Alice's Adventures Underground*, the Duchess is indeed absent, but many traits of her character are attributed to the Queen of Hearts.[206] So, how many characters are there in situations in which fusion or fission take place?

The examples above show that we need a well-defined criterion for determining when a literary character remains the same in the passage from one work to another, even if it sometimes changes its features radically in the different stories. My work is intended to provide such a criterion. I do not aim to develop a semantic analysis of fictional and metafictional sentences in order to establish whether or not they are true and under which conditions. Nor I intend to undertake a metaphysical investigation to answer the question whether or not these sentences carry ontological commitment to fictional and mythological entities. Instead, I shall try to deal with issues concerning the cognitive nature of fictional characters and the way in which we conceive them into our mind. I think that, by properly using the theory of mental files, we can account for all the puzzles we mentioned above.

[206] The example is suggested by Voltolini (2010: 83).

3.6 Identity at the fictional level

First of all, I shall consider the identity of literary characters inside a single work of fiction. I argued that the generation of files in fictional contexts works as in everyday life. When the author introduces new characters, we open new files storing information that the story tells us: their physical appearance, sex, age, actions and any other kind of properties. All files created by reading a novel are indexed to that world of fiction: whenever it is possible, they are linked to our regular files, otherwise we have a free-wheeling use. Regardless of whether they are loaded or unloaded, the structure of files is the same. We label the file with the name of the character it aims to refer to. So, for instance, we come to have in our mind the files <SHERLOCK HOLMES, *A Study in Scarlet*>, <NAPOLEON, *War and Peace*>, and so on.

So far, so good. But things are not always so simple. Sometimes we do not know how a certain character is called. The name may remain unknown because the author wishes to leave a mystery about one individual and the absence of the proper name is an integral part of the story. A famous example is the *Innominato* in the *Promessi Sposi*. This case is fairly straightforward. We can open a mental file and use some definite descriptions as a label to identify the character[207], such as <THE NAMELESS, *Promessi Sposi*>. Often a character does not have a name because its role in the plot of the story is not important enough. It is, so to say, a mere background actor. Let us consider this situation: we are walking down the street and meet several strangers passing by: we do not know their names but nonetheless we may identify people around us by using indexicals or demonstratives, like "that guy in front of me", or definite description like "the tall man with the hat". Similarly, we can use indexicals, demonstrative or definite description to individuate secondary characters in novels, for instance: "The man the protagonist was talking with in the first chapter of the book".

An interesting case of unnamed character is presented in *À la Recherche du Temps Perdu*. Here, the protagonist of the story and the narrator overlap: the book is a sort of

[207] The description is used referentially, and not attributively. We have a referential description when it is used only to pick out an object, but the referent is determined by an ER relation that already exists, as in Donnellan's example "the man drinking a martini". See ch. 2.7.

autobiography, in which a first-person voice tells us the protagonist's recollections of experiences from childhood to adulthood. Suppose that, in the whole *Recherche*, his name is never told. This does not represent a problem: for our purposes, it is sufficient to open a file like <THE NARRATOR, *À la Recherche du Temps Perdu*>. Such case, once again, helps us to clarify the usefulness of having indexed files, since the narrator of the *Recherche* must not be confused with its author. Even if a few details in the book are directly inspired by Marcel Proust's real life, the *Recherche* is a fictional novel, not a trustworthy biography: many events, characters and places are invented or freely gathered from reality. The use of the first person does not by itself allow us to overlap the real Marcel Proust with the fictional storyteller of the novel[208].

As we have seen, in the vast majority of novels it is easy to individuate fictional characters, regardless of whether we know their names or not, just as in ordinary life it is easy to recognize objects around us. But, just as in real life there are exceptional cases that originate misunderstandings, so there are in fictions. A subject, without any astronomical background, may open two separate files HESPERUS and PHOSPHORUS for the same planet Venus. Once the mistake is recognized, i.e., that the names "Hesperus" and "Phosphorus" refer to a single celestial body observed in two different moments, the subject links the two files. Situations like this have already been widely commented by philosophers[209]. Let us now analyze a literary variant of the Hesperus/Phosphorus case: the *Strange Case of Dr. Jekyll and Mr. Hyde*. In the novel, a character is depicted as two distinct ones: sometimes he appears as the gentle and kind Dr. Jekyll, sometimes as the violent and brutal Mr. Hyde. Only at the end of the reading we learn that Dr. Jekyll and Mr. Hyde are alternative personalities of the same fictional character. Misled by the narration of the events, we open, at first, two files, one for Dr. Jekyll and one for Mr. Hyde. Later, we link the two files, even if the characters of the novel seem very distant from each other. But which kind of connection are we using? According to Recanati (2013: 155), in addition to the two forms of linking that we have already presented – horizontal linking between regular files and vertical linking between regular and indexed files[210] – indexed files require the introduction of some other forms of linking. The one that matters for our case is *internal linking*, that represents a

[208] At this regard, see Bonomi (1994: 14-16).

[209] See ch. 2.2 and 2.3.

[210] See ch. 3.2.

connection existing only in the mind of another subject, or in the speaker's mind at a certain time in the past:

> Internal linking reflects *the subject's belief in some identity*, whether the subject is the speaker/thinker or some other subject whose point of view the speaker/thinker is representing. It is only in the case of internal linking that is it possible to represent linking by entering identity information into the linked file. (Recanati 2012: 191)

To illustrate this point, Recanati proposes the following example. Mistakenly, Paul used to believe that there were two distinct people, Bert and Tom, while in fact there is one single person. Now he has discovered the truth, so he believes that Bert is Tom. But we, who are skeptic about the identity, report his doxastic state by saying:

11) Paul believes that Bert is Tom

The two files respectively associated with the names “Bert” and “Tom” are indexed to Paul, since it is Paul that accepts the identity “Bert = Tom”. They are represented as linked in Paul's mind, not in ours. We cannot use horizontal linking, because it is a connection that works only for regular files, and our regular files are not linked. Nor vertical linking can be useful for this task, for we are not connecting a vicarious file with a regular one. The case posed by Dr. Jekyll and Mr. Hyde is similar. We, as readers, do not know the identity between the two characters until we get to the end of the story, in which it is told that Jekyll and Hyde are the same person.

We can conclude that even when we know the names of all the characters, we may be doubtful about how many characters there are[211]. In the reality, empirical investigations determine whether two names refer to the same object. On the contrary, within the fiction we completely defer to the author's choice. It happens as if we come to believe that “Bert” and “Tom” refer to a single individual just because Paul thinks they do. The situation is not so bizarre, at a closer insight. In fact, we really do not have regular files on Dr. Jekyll and Mr.

[211] Other interesting literary cases are twin, look-alike, double or doppelganger characters. Here the mechanism is the opposite: we start with the belief that there is a single character and then we discover that they are two or more.

Hyde: the only files in our mind are those indexed to the story. Therefore, I agree with Tim Crane:

> But there is a very important additional fact about the fiction: the author's stipulation that they are nonetheless identical. This is itself a representation with a very special role of trumping all these differences in representations. [...] Representing things as identical is the ultimate way of representing them as similar, despite other dissimilarities. But this has to be something claimed in the story. No sense can be made of the idea that two characters in a story might "really" be one, if the author of the story does not say so. (Crane 2013: 167)

Thus, sentence:

8) Jekyll is Mr. Hyde

is true in the world of the fiction: despite the fact that Dr. Jekyll and Mr. Hyde are described as two distinct people, with different and almost opposite features, they are one single individual because of the author's stipulation. Once we link the files, information can be shared. So the file <DR. JEKYLL, *Strange Case of Dr. Jekyll and Mr. Hyde*> contains the information 'being Mr. Hyde', and the file <MR. HYDE, *Strange Case of Dr. Jekyll and Mr. Hyde*> contains the information 'being Jekyll', as it is explicitly stated in the story. Precisely because it is claimed inside the novel, the identity between Jekyll and Mr. Hyde, expressed by (8), is a fact concerning the fiction, that we must accept as true.

3.7 Identity at the metafictional level

Problems of identity arise not only within a single story but also in the comparison between two (or more) works of fiction. These cases are much harder to treat, since there is not a thing like the authority of one single writer to which we can appeal. Come back to sentence:

9) Homer's Odysseus and Dante's Ulysses are the same fictional character

As we said, (9) generates conflicting if not even opposed intuitions. But what do we really mean with it? When we claim that Homer's Odysseus and Dante's Ulysses are the same fictional character we are saying that the relevant representations aim to refer to the same individual, in virtue of the fact that the latter character is causally connected to the first one or, in other words, that Dante, in the creation of his Ulysses, was inspired by the character featured in the Greek epic poems. At a cognitive level, we can translate the idea by saying that the corresponding mental files are vertically linked to each other. In fact, when this situation occurs, the latter file inherits the referent of the former, *so far as reference is possible*. In this case, since both files are vicarious and unloaded with respect to a regular one, they do not refer, but still a unique referent is presupposed. They would have the same referent, if there was one. Moreover, due to the vertical form of the link, the content of each file is preserved. We do not mix information deriving from separate fictions. In Homer's poem, Ulysses, after a journey lasted ten years, comes back to Ithaca and restores his reign, whereas, in the *Inferno*, Ulysses tells Dante that he set out with his men from Circe's island for a journey of exploration beyond the Pillars of Hercules and then he died after a shipwreck. We are not surprised by these inconsistencies, because we look at the two stories as two alternative versions of the myth of Ulysses. Although we can import some pieces of information from the Greek poems into the *Divina Commedia*, in order to enrich our comprehension and appreciation of Dante's work, importation is not automatic: it only takes place as far as the *Divina Commedia* allows us so.

Before going on, we need to focus further on vicarious files. As we know, they are characterized by an indexed structure[212]. What we need to add is that their structure can also be recursive:

> Indexed files are recursive: the file component of an indexed file may itself be an indexed file. Thus S_1 may think about S_2's way of thinking of S_3's way of thinking of some entity, and to that effect may entertain the indexed file <<*f*, S_3>, S_2>. (Recanati 2012: 183)

This property is very useful for us. In fact, I suggest that linking between vicarious files can be represented by means of their indexed and recursive structure. For the sake of argument, suppose that Dante was inspired for his Ulysses solely by the *Odyssey*. The original file is <ULYSSES, *Odyssey*>. Now, the *Divina Commedia* depicts a personal interpretation of Homer's Ulysses. Thus, according to Dante's fiction, we are provided with an alternative way of imagining the hypothetical referent of the file <ULYSSES, *Odyssey*>. We obtain a recursive file <<ULYSSES, *Odyssey*> *Divina Commedia*>.[213]

I maintain that, in virtue of their linking, vicarious files gather into networks, whose individual knot is given by a file indexed to a fiction in which the character in question appears. Other authors have already used the notion of network, but with different meanings and purposes. Perry (2001a), Everett (2013) and Friend (2011, 2014), in fact, use the term "network" to explain the phenomenon of co-identification in the case of empty names and to give an account of how more people can share the same mental representation. These issues are beyond the scope of my work. It should be emphasized, however, that these authors conceive networks as sets of relations between regular files. Instead, in my perspective, there are not regular files associated with empty[214], but only vicarious files indexed to fictions. Networks develop at the level of vicarious files.

By relying on the notion of network, we can justify all the different intuitions on the identity of fictional characters that we exposed above. On one side, we provide an account for the idea that there is an Ulysses belonging to the *Iliad*, one to the *Odyssey*, another to the

[212] See ch. 3.2.

[213] The original file tells us how to represent Ulysses according to the *Odyssey*, the second one how to represent, according to the *Divina Commedia*, the Ulysses depicted in the Odyssey.

[214] I put in doubt this idea in 3.10.

Divina Commedia, and so on, by saying that each version of Ulysses corresponds, in our mind, to a specific vicarious file indexed to the relevant fiction. On the other side, we can also explain the intuition that the same fictional character persists in the transition from one work to another, a character that is Ulysses and not just "the Ulysses of some fiction". The continuity between the various literary characters is given by the presence of a unique network that links all the files together. All these files, vertically linked to each other, aim to refer to the same individual and they would do, if such individual existed.

Two fictional characters count as variants of a single character when, at the cognitive level, the relevant indexed files are located within the same network. The notion of network allows us to include any new representation of Ulysses. But we need a clear criterion to determine whether a file belongs to a network or not: a literary character counts as an alternative version, or remake, of another literary character already existing, for example Ulysses, if and only if its author, by creating the character, takes inspiration from a preexisting work on Ulysses. In other words, if there is a causal connection between the two stories. When this happens, it means that the relevant file fits in the Ulysses network. Otherwise, a new network of files is generated.

By means of the criterion just mentioned and the recursive structure of indexed files we can also explain in which ways network develops. Consider the example of paladin Roland. His epic adventures and death are told for the first time in *La Chanson de Roland*. By this old French poem of the 11th century was later inspired Boiardo, an Italian poet of the Renaissance. Boiardo wrote the *Orlando Innamorato*, whose Ariosto's *Orlando Furioso* is a further continuation. The network of mental files has here a linear expansion from *La Chanson de Roland* to the *Orlando Furioso:*

a) <ROLAND, *Chanson de Roland*>

b) <<ROLAND, *Chanson de Roland*> *Orlando Innamorato*>

c) <<<ROLAND, *Chanson de Roland*> *Orlando Innamorato*> *Orlando Furioso*>

We can now consider the more complicated case of Ulysses. Let us say, even if we simplify the story, that the character appeared for the first time in the *Iliad*, then in the *Odyssey*, and finally in the Latin *Aeneid* by Virgil. After centuries Dante, who could not read

the Greek poems, took inspiration from Virgil and from other Latin sources[215] for the damned soul appearing in the *Divina Commedia*. Finally, James Joyce wrote his *Ulysses*, but taking as a model for Leopold Bloom not the figure of Ulysses of the *Divina Commedia*, but the one of the *Odyssey.* The network has, at first, a linear development:

a) <ULYSSES, *Iliad*>

b) <<ULYSSES, *Iliad*> *Odyssey*>

c) <<<ULYSSES, *Iliad*> *Odyssey*> *Aeneid*>

At this point, it splits into two branches: in one branch we have Dante's Ulysses, linked to Virgil's poem:

d) <<<<ULYSSES, *Iliad*> *Odyssey*> *Aeneid*> *Divina Commedia*>

In the other branch, we have Joyce's novel, which is inspired by the *Odyssey*. Leopold Bloom is at the same level as Virgil's Ulysses, for both are vertically linked to the same file:

c') <<<ULYSSES, *Iliad*> *Odyssey*> *Ulysses* $_{\text{Joyce}}$ >

It is now time to move on to the case in which a file cannot be part of a preexisting network, therefore giving rise to a new fictional character. I borrow the example from Voltolini:

> Take Jorge Luis Borges's famous story of a man called Pierre Menard who happens to write a text that is word for word identical with Miguel Cervantes's *Don Quixote*. Assume, in this variant of Borges's story, that Menard and Cervantes are unknown to each other, even though they live in the very same town; one can even suppose that they are neighbors. In that case, the Borges story describes a situation in which one and the same set of properties

[215] Another important source of information for Dante was Ovid's *Metamorphoses*, but I leave Ovid aside from the analysis.

> corresponds to different fictional objects: Cervantes's Don Quixote and Pierre Menard's Don Quixote are two distinct fictional characters who, nonetheless, share all the properties they have in the respective works. (Voltolini 2006: 33)

The puzzle is usually presented as an objection to Meinongian and neo-Meinongian theories[216]. In fact, given that Cervantes' and Menard's Don Quixote share the same properties, according to (neo-)Meinongian theories they should be a single character, whereas the goal is precisely to distinguish them in some way. My view does not suffer from this objection. Under the assumption that Pierre Menard does not mean to refer to any other work, we cannot qualify his character as a new version of Don Quixote, no matter how many properties they have in common. Menard's Don Quixote is not linked to the network of files that is originated by <DON QUIXOTE, *Don Quijote de la Mancha*>. It is, indeed, an entirely new Don Quixote, which at best will produce an alternative network of indexed mental files.

Finally, my theory can also account for phenomena of fission and fusion. I begin with the fusion case. We have a fusion when an author intends to import into a work of fiction *W'* two characters *x* and *y* from a previous work *W* as a single character *z*[217]. Let us recall the example:

> In a version of the *Recherche* dated 1912, two characters appear, the musician Berget and the naturalist Vington, while in the final version we only can find the musician Vinteuil, who somehow summarizes Berget's and Vington's attributes. Now we can say that Proust, obviously well aware of Berget and Vington elaborated in the draft of 1912, intends to import in the final version of the *Recherche* both Berget as Vinteuil and Vington as Vinteuil. (Voltolini 2010: 83, *my translation*)

To account for this case, it is sufficient to observe that since Vinteuil is inspired to both Berget and Vington, these connections must also be reflected at the level of mental files. In other words, the file on Vinteuil is vertically linked to both Berget's and Vington's files:

a) Berget: <BERGET, *Recherche* $_{1912}$>

[216] See ch. 1.4.3. See also Voltolini (2006: 32-35) and Voltolini & Kroon (2016).

[217] See Voltolini & Kroon (2016) and Voltolini (2010).

Vington: <VINGTON, *Recherche* $_{1912}$>

b) Vinteuil: <<BERGET, *Recherche* $_{1912}$>;<VINGTON, *Recherche* $_{1912}$>*Recherche* $_{1927}$>

In fact, there is no reason to suppose that a file can have only a single vertical link and a structure of the type <*f*, S>. Why not to admit also files that have a structure like <<f_1, f_2> S>? This idea accounts for the fact that an author, by creating a character, can take as a model two or more characters already imagined by himself or by others. One could write, for example, a tear-jerking story whose heroine has attributes drawn from both Flaubert's *Madame Bovary* and Tolstoy's *Anna Karenina*. By this way, we admit the possibility that different networks of mental files come, at a certain point, to overlap and have common knots. This constitutes a further advantage for my proposal: in fact, it accommodates cases like that of Santa Claus and Father Christmas, discussed by Salis (2013) and Friend (2014). Originally these names were used to indicate characters belonging to two different folk traditions; then people began to identify them as a single character: at this point, a merger took place between networks of indexed files that were initially independent.

Let us move now to an example of fission of literary characters, like the one in Carroll's *Alice in Wonderland*, where the Duchess and the Queen of Hearts derive from the splitting of a previous Queen of Hearts featured in the draft *Alice's Adventures Underground*. The explanation is easy. Since the two new characters are inspired by the same model, at the level of mental files we have two vicarious files, indexed to *Alice in Wonderland*, both vertically linked to the file indexed to *Alice's Adventures Underground*. The network splits into separate branches, in a way not dissimilar from what we have seen in Ulysses' case:

a) original Queen of Hearts*:* <QUEEN OF HEARTS, *Alice's Adventures Underground*>

b) the Duchess*:* <<QUEEN OF HEARTS, *Alice's Adventures Underground*> *Alice in Wonderland*>

b') new Queen of Hearts: <<QUEEN OF HEARTS, *Alice's Adventures Underground*> *Alice in Wonderland*>

3.8 Mixed cases: conniving and non-conniving perspectives compared

We discussed about literary characters both at the fictional and at the metafictional level. As for the fictional level, we defer to the author's authority to determine the identity of literary characters. We said that two vicarious files are internally linked when they refer to the same character according to a certain work of fiction. As regards the metafictional level, we proposed the notion of network between indexed files to account for identity problems in the passage from one text to another or from an earlier version to a latter one of the same novel, with related phenomena of fusion or fission. We can now recall the dispute between Cervantes and Alonso Fernández de Avellaneda and see how to address this ambiguous case. Coherently with the position I have expressed, I argue that, at a metafictional level, Avellaneda's Don Quixote is as legitimate as Cervantes's one, since both the authors took as a model the story of 1605, in which the character of Don Quixote appeared for the first time. In terms of mental files, we have:

a) <DON QUIXOTE, *Don Quixote* $_{1605}$>

b) <<DON QUIXOTE, *Don Quixote* $_{1605}$> Avellaneda's *Second Part* $_{1614}$>

<<DON QUIXOTE, *Don Quixote* $_{1605}$> Cervantes's *Second Part* $_{1615}$>

The latter two files are vertically linked to <DON QUIXOTE, *Don Quixote* $_{1605}$> and are part of the same network of mental files. Therefore, it is not a case like Pierre Menard's. But this fact does not clash with what Cervantes claims in the *Preface* of 1615. Here, Evans's distinction between conniving and non-conniving uses plays a crucial role. In fact, sentence:

10) Avellaneda's Don Quixote and Cervantes's Don Quixote are not the same character

may be either true or false. In its conniving use, as a fictional sentence uttered inside the fiction created by Cervantes, (10) is true because the author himself states it. But (10) is false, according to my perspective, if we interpret it in a non-conniving way, as a metafictional sentence.

These considerations allow me to say that identity (or non-identity) between two characters, as established by an author within a novel, does not necessarily reflect the situation at the level of network of indexed files. Let us recall our previous analysis on Dr. Jekyll and Mr. Hyde:

8) Jekyll is Mr. Hyde

Even if (8) is true inside the story, it does not guarantee that the relevant indexed files are actually vertically linked[218]. There may be a discrepancy between the identity of the characters in the novel and the network of files at a cognitive level. We can say that Dr. Jekyll and Mr. Hyde are two different characters that correspond to two distinct representations or vicarious files. The figure of Hyde, after all, is more fascinating than Jekyll and certainly had more success than its good "twin". Why should we take them as a single character, when we can claim that Jekyll and Mr. Hyde are two distinct characters that in the novel coincide with one?[219] Imagine writing a remake of Stevenson's story in which, by changing only the final, it is claimed that Jekyll is not Mr. Hyde. Is it a case of fission such as the one discussed in the previous session? I am tempted to give a negative answer. Moreover, the identity of Jekyll and Mr. Hyde, precisely because it is claimed inside the novel, appears to involve a fact concerning the fiction, and not a metafictional truth. As in Don Quixote's example, sentence expressed by (8) with a conniving use is not the same sentence expressed by (8) with a non-conniving use: in the first case, one utters a fictional sentence, in the second a metafictional one, and these sentences have different truth conditions since the context of evaluation changes.

[218] Probably they are, but what I want to claim is that internal linking does not merge files. We still keep separate Jekyll's and Hyde's files.

[219] Suppose that there was a network for Jekyll and a network for Hyde, and then Stevenson decided to import the two characters as a single one in his novel. We would deal with a phenomenon of fusion. Superman and Clark Kent is a comparable example. See Salis (2013) and Friend (2014).

3.9 Reply to a possible objection: the general character

I would like to consider a possible objection that can be moved to the analysis so far developed. The objection is that, according to my theory, we can think about a fictional character only vicariously, passing through the way in which the character is depicted by some work of fiction. One may wonder whether this condition is too much restrictive. I report a passage from Voltolini:

> To the extent that they do not share the same properties assigned to them in the course of the relevant narratives, the Roland of *Orlando Innamorato* is not [...] the Roland of *Orlando Furioso*. Of course, [...] all these fictional characters can be linked to a sort of *general* character: the Roland in general, the Faust in general, the Ulysses in general. But it is not clear how this general character should be featured.
>
> (Voltolini 2010: 77-78, *my translation*)

On the one hand, it seems that we can think about literary characters even independently of the works of fiction in which they appear. On the other hand, it is hard to explain how characters may be conceived outside their relative stories. At a closer look, the issue can be divided into two questions. First question: do we really need to use vicarious files to think about literary characters, thereby appealing to specific fictions? Second question: is there an Ulysses in general, to which all the other are linked?

Let us begin with the first question. A suggestive hypothesis that deserves to be explored is the following: Recanati suggests that, in addition to vicarious files, we can think about fictional characters, such as Ulysses, also by means of regular files. In this case, we explicitly violate the acquaintance constraint through a mechanism of simulation. By doing so, we represent the fictional character as if we had direct access to it, without deploying the filing system of other subjects. Our thought is not globally meta-representational, as it happens with free-wheeling indexed files. Then, it is plausible to say that, since we are not

representing someone else's way of thinking about Ulysses, i.e. Ulysses according to a particular story, the pretended reference of our primary file is the Ulysses in general. This perspective is reinforced by the parallelism that we could trace with what happens for non-empty names. When we read Shakespeare's play dedicated to Julius Caesar, we open an indexed file where we store information deriving from the fiction, but the file does not remain unloaded for we already have a regular file to link to. We can import notions from the regular to the vicarious file and, potentially, export some pieces of information. A similar situation could also emerge for empty names. We could simulate reference with the general character of Ulysses by deploying a regular file and then, every time we read a story about Ulysses, generate specific vicarious files vertically linked with our regular one, allowing importation and exportation of fictional information. The reason why I reject this interpretation is that I give a negative answer to the second question we asked before, i.e., I do not believe that it is possible to conceive literary characters independently from any work of fiction.

I believe that it is idle to look for a file with the label "*The Ulysses in general*" or similar ones. Suppose that this file is a regular one. It would be either too broad or too tight. It would be too broad in case it contained all the information stored in vicarious files, because such information would generate inconsistencies or contradictions.[220] In the *Odyssey*, Ulysses comes back to Ithaca after ten years of travels, in the *Divina Commedia* he died without having ever returned home. The two versions of the myth do not coincide. We have two options: either the Ulysses in general is a set of contradictory attributes, or there is no general character conceived in this way. The file would be too tight if it should contain only the "essential" data on an *Ur*-Ulysses, i.e. information that delimits a set of unavoidable and non-contradictory properties[221]. First of all, the proposal fails because we cannot account for Pierre Menard's case, a puzzle that I take to be a very serious obstacle. Secondly, because each author is free to interpret a fictional character as she likes and to chose the set of properties she prefers to create a new version of it. Any limitation would be arbitrary. One can, for example, write a novel on Ulysses, taking inspiration from Homer's poems, in which the protagonist is not a Greek hero. And, in fact, we have the *Tale of Yuriwaka*, a Japanese rendition of the *Odyssey,* introduced in the country by Christian missionaries. Or there can be a story in which Ulysses never traveled by sea: Joyce's novel is indeed a case like this. With

[220] For a discussion on this topic, see Reicher (1995).

[221] I thank Carola Barbero for the discussion we had on this issue. For some proposals in favor of an essentialist view of fictional characters, see Eco (1979, 2000) and Orilia (2002, ch.14).

parodies, as Guido Gozzano's *Ulysses is shipwrecked ... on board of a yacht*?, differences with the original myth can be even bigger.

One can argue that if a character dissociates too much from the canonical model, it can no longer be considered a legitimate version of Ulysses. But it is hard to find sharp and objective criteria to define the canonical model. Which properties must be taken as relevant in order for a character to be Ulysses? Any choice would exclude some versions of Ulysses that we would rather include. Of course, in order for a new Ulysses to be recognized as such, there must be some similarities between it and a preexisting version of Ulysses. But the set of properties shared by the new Ulysses and the old one is contingent, there are no essential attributes that one needs to keep. Dante's Ulysses has some similarities with Virgil's one, Virgil's Ulysses has some similarities with Homer's one, but nothing guarantees that Dante's and Homer's Ulysses resemble each other. Imagine comparing all the stories of Ulysses provided by the whole literature. We could arrive at the situation in which between two versions of Ulysses there are no shared properties. And indeed, which properties are in common between Homer's, Virgil's, Ovid's, Dante's, Tennyson's, Joyce's, Saba's, Walcott's and Gozzano's characters?

It is possible to overcome these problems by proposing a cluster theory, according to which in order for a character to be considered Ulysses, it must possess not all the essential attributes of the model, but a sufficient amount of them[222]. The model, or general character, is thus not composed by a single list of essential attributes, but by a cluster of attributes among which an author can choose. This theory still needs to answer to Pierre Menard's puzzle. Moreover, difficulties remain in determining which attributes are relevant enough to be part of the model and which number of them must be included in the creation of a new Ulysses.

It could be assumed that the *real* Ulysses must be individuated, at the level of vicarious files, in the first file of the network, to which all other indexed files are directly or indirectly vertically linked. This position does not seem so good. In fact, nothing guarantees that the first work in which a literary character appears is also the most culturally significant and cognitively relevant. For example, in the creation of *Dracula*, Bram Stoker was inspired by John Polidori's *The Vampire*. Thus, our indexed files about Stoker's character has the structure <<VAMPIRE, *The Vampire*> *Dracula*>. Nevertheless, we do not refer to Polidori's novel, in most cases, when we talk or think about vampires, but to the story written by Bram

[222]I thank Tommaso Piazza for this suggestion.

Stoker. Our mental file <<VAMPIRE, *The Vampire*> *Dracula*> has more cognitive importance than <VAMPIRE, *The Vampire*>, to which the first is linked. Similarly, many people who never read the French poem will give a greater cognitive weight to the Roland depicted by Boiardo or by Ariosto rather than to the one featured in the original medieval work. Finally, even if the *Iliad* precedes the composition of the *Odyssey*, according to the most accepted chronology, it is the traveler of the latter that we all have in mind, not the warrior of the former. It makes little sense to wonder which one is the real Ulysses, as it makes little sense to ask which is the real Roland, whether the one in *La Chanson de Roland*, or the one in *Orlando Innamorato*, or the one in *Orlando Furioso*. Each of us will grant his preference to a certain Roland, and maybe will have in mind, for example, the specific figure described by Ariosto rather than the character in the *Orlando Innamorato*, or in the *Chanson de Roland*, but this does not mean that the file based on the *Orlando Furioso* has a general value, that it is the file of the real Roland.

So, what does exist? It exists the continuity from one fiction on Ulysses to another, the chain of causal relations that generates a network by mean of which indexed files are embedded in our minds. From this network we should not expect to derive a single overall concept – an essence – of Ulysses. Intuition according to which we can think about literary characters even apart from the single stories in which they appear is misleading. It is true that we do not focus our attention on specific works of fiction whenever we think about a character. However, our mental representation is always based on some particular fiction, the one that is most relevant for us. We know that Ulysses is a cunning, brave and ingenious man. He is an eloquent speaker and a valent warrior. He fought in the Trojan war, sailed for ten years before returning home, defeated the cyclops Polyphemus, listened to the Sirens' song, and so on. This is exactly the image we get from Homer's poems. One can object that the *Iliad* does not tell us anything about the stratagem of the Trojan Horse, while the *Odyssey* mentions it only marginally, but nonetheless it represents one of most famous episodes in Ulysses's life. The most detailed and familiar narration of the Trojan Horse is in Book II of Virgil's *Aeneid*, which was written several centuries later the Greek poems. How can we explain this fact? The point is that the story of Ulysses was not the creation of a single author, but it derives from a series of myths and folktales that were told by anonymous bards in ancient Greece. It is from those stories that *Iliad* and *Odyssey* originated and, with them, all the other Greek and Latin poems. There is nothing problematic about this. Some fictional characters – like Pinocchio, Sherlock Holmes, Anna Karenina, Emma Bovary, Don Quixote –

have copyright, in the sense that there is an author who created them and one or more works of fiction that represent the canonical source of information. Some other characters – like Ulysses, Roland, Vampires, Father Christmas – belong to the folk or mythological culture of some countries. In this case, the source of information is vaguer and attributes that are considered standard for a fictional character may change through time. There may be, at a certain point, a literary work that somehow fixes the features of a character or adds new ones, but the most prominent source of information remains the undefined heritage of stories and legends from which it originates. We can store such information into a vicarious file that is indexed to legends and folk tales, like <ULYSSES, *Greek myth*>. This file is at the beginning of the network and it is probably the one that we use to represent the Ulysses in general. But, still, this Ulysses is not imagined independently from any fiction, but in accordance to a vague and indeterminate fiction, the one provided by the myth.

3.10 Regular files

So far, I have discussed about vicarious files and the role that they have in our understanding of fictional characters. It is now time to turn our attention also to regular files. As I said, I disagree with Recanati in the idea that regular files can be used to simulate a direct or mediate relation of acquaintance with a character so as described by a fictional story. Does this mean that we must exclude regular files from the account? I am tempted to give a negative answer. Before going on, let me explain my point of view. I began my analysis by considering the use of real names in fictional stories and I argued that they work in the same way as names of purely fictional characters do. When we deal with a novel, both real and fictional names lead us to open indexed files in which we store information derived from the novel. But similarities stop at this point. In fact, in the case of real names the vicarious file is vertically linked with the regular file that we already possess. On the contrary, in the case of fictional names we have a free-wheeling, or unloaded, use of indexed file. However, I think that we can push the analogy further and make room for regular files even in the latter case. The

reason why we should do so is that there are two different ways we can speak about literary characters: we can say something, for instance, about Madame Bovary either from the inside or from the outside of the novel itself. A sentence like:

12) Emma Bovary falls in love with Rodolphe Boulanger

is a fictional sentence, for it does not involve any genuine fact about the real world, but a fact related to the fiction. It has truth-values only insofar we are engaged in a practice of make-believe. On the contrary, sentence:

13) Emma Bovary is a fictional character

is a metafictional sentence that predicates something about a literary character from a perspective that is external to the fiction the character belongs to[223]. In fact, Madame Bovary truly is a fictional character according to our world. So, we have here two kinds of information corresponding to two distinct ways we can speak about a literary character: some pieces of information concern the fiction; others do not. Due to this reason, I think that we should not treat all the information on Bovary as being of the same sort, i.e. as belonging to a single file. Vicarious files only contain information drawn from the story. But it is not true, in Flaubert's novel, that Emma Bovary is a fictional character. On the contrary, she is depicted as a flesh-and-blood woman. Metafictional information is derived from our comprehension of literature and our knowledge of the world. As informed readers, we know that literary works are usually full of characters that are entirely made up. We also know that Emma Bovary never existed and so any chronicle of her romantic adventures must be a figment of imagination.

There is another important difference between fictional and metafictional sentences. Generally[224], the former refer to literary characters as concrete individuals that exist

[223] For the rest of the chapter, for "metafictional sentences" I will only refer to external metafictional sentences. For a discussion on how to apply mental files to internal metafictional sentences, see Recanati (2018).

[224] There are some exceptions due to the existence of "metafictional" narratives. See ch. 1.3.1 and Voltolinini & Kroon (2016) and Voltolini (2006: 214-215).

spatiotemporally and do actions like living, thinking, having feelings, dying. On the contrary, metafictional sentences refer to literary characters as abstract objects, i.e. something that does not exist spatiotemporally and can be created by some author, can move from a fictional work to another, has aesthetic and literary values, and so on[225]. As Recanati observes:

> We have seen that fictional characters are two-sided entities; or rather, that the concept of a fictional character is a two-sided concept [...] We construe fictional characters as both flesh and blood individuals (viewed from inside the pretence) and abstract artifacts (viewed from outside). This is not incoherent, since each particular predication selects the aspect it applies to: sometimes we talk about the flesh and blood individual, and sometimes about the abstract artifact. (Recanati 2018: 24)[226]

One of the most difficult challenges for any theory of fiction is indeed to account for this two-fold nature of literary characters. But one point is clear: if we stored fictional and metafictional information in the same file, we would keep together incompatible pieces of data, thus generating inconsistencies and even contradiction. Emma Bovary cannot be, at the same time, a concrete individual such as a flesh-and-blood woman and an abstract entity such as a fictional character. Moreover, consider sentence:

14) Sherlock Holmes is famous all over the world.

Of course, it is true that Sherlock Holmes, as a fictional character, is famous: indeed, it is one of the most famous characters of the whole literature. But inside the novels written by Conan Doyle, Sherlock Holmes is a quite unknown detective. So in a single file we should store two contradictory predicates: “being famous” and “not being famous”.

For all these reasons, I think that we need to match the vicarious file with a regular file. The former collects information about Emma Bovary drawn from Flaubert's novel, for

[225] Here, I endorse an artifactualist account, for it seems to me the best one for dealing with fictional objects when we talk about them from a metafictional perspective, but I cannot tackle in detail this topic. However, the theory of mental file is also compatible with the neo-Meinongian framework. So one can remain neutral about the nature of fictional entities, whether they must be conceived as eternal and immutable Platonic objects or created and contingent artifacts.

[226] See also Terrone (forthcoming) “Twofileness: Fictional Characters as Generators of Mental Files”.

instance “having adulterous affairs”, “committing suicide”, and so on. We pretend to attribute this kind of information to a real woman whose tragic life is narrated in the book. It is this file that we use when we deal with the fictional sentence (12). If in the file <BOVARY, *Emma Bovary*> there is the predicate “falling in love with Rodolphe Boulanger”, we consider (12) as true, since it is what the story says. On the contrary, the regular file BOVARY contains information from a perspective external to the context of pretense, i.e. information about Bovary as figment of an author's imagination such as “being a fictional character”, “being invented by Gustave Flaubert”, “being the main character of the novel”, and so on. The regular file is based on, or simulates, a relation to an abstract object[227], which is no more the woman Bovary. When we consider metafictional sentences like (13), we use such kind of file.

I suggest that we need regular files for storing metafictional information because, at this level, we are talking about fictional characters as abstract objects, without any engagement in games of make-believe, so it does not make sense to use vicarious files indexed to some fiction. Moreover, when we utter a metafictional sentence, we are not reporting someone else's point of view, by putting ourselves in another subjet's shoes, but we are expressing our own believes, with the aim to say something that is genuinely true or false about literary characters, exactly as we do when we use our regular files about Julius Caesar or Napoleon to express what we know about those people, not what Shakespeare or Tolstoj claim in their stories. I do not want to go in details on which type of relation we have with abstract objects, or whether there are or there are not those kinds of entities. It is sufficient to say that reference, or pretended reference, to abstract objects normally takes place in almost every kind of discourse, when we talk about numbers, laws, classes, values, sports and literary works themselves. To account for these uses is a general problem for the theory of mental files that I do not have the possibility to address in this work[228].

Once we admit the idea to have both regular and vicarious files for literary characters, we can say that the latter are vertically connected with the former. This linking accounts for the two-fold nature of literary characters: when we talk about the concrete individual, we deploy the vicarious file; when we talk about the abstract object, we deploy the regular file. Since files are linked, sameness of reference is maintained. But regular files also allow us to

[227] This point depends on the ontological point of view that you adopt. If you are a realist, you would admit that we have some sort of relation with abstract objects, so there is no simulation. If you are an antirealist, abstract objects are the products of a simulation. I remain neutral on this issue.

[228] Some cues are provided by Recanati (2018) and Terrone (2017).

say something more about the phenomenon of migration of fictional characters that we have already analyzed before. The fact that all vicarious files, corresponding to specific versions of a single character appearing in different works of fiction, are linked to one regular file, reinforces the impression that we deal with the same character that moves from a work of fiction to another. We do not need to postulate a general fictional character, i.e. the Ulysses in general, collecting all the essential attributes of the Greek hero, but an abstract object depicted from time to time in different ways depending on the fictional works we are considering[229].

Depending on the situation, we can use the regular or the vicarious file to interpret a sentence as either fictional or metafictional. Consider the following example, taken from Voltolini:

15) The Father is a fictional character

"Being a fictional character" is a typical predicate to be put into the regular file. However, there are metafictional narratives whose protagonists are not characterized in the usual way as flesh-and-blood individuals, but instead as fictional characters. The Father in Pirandello's *Six Characters in Search of an Author* is the case in point[230]. So the predicate "being a fictional character" can be assigned to both a regular or a vicarious file. But in one case we obtain a metafictional sentence, since we are using information contained in the file FATHER, in the other case a fictional sentence, derived from the file <FATHER, *Six Characters in Search of an Author*>. Thus, my claim is that, at the cognitive level, the difference between a fictional sentence and a metafictional sentence, or (to recall Evans' distinction) between conniving and non-conniving uses of a fictional sentence, is explained by the type of files that we deploy. The same analysis also applies for the example of Sherlock Holmes we mentioned before. Sentence (14) can be read in two ways and, depending on whether we consider (14) a fictional or metafictional sentence, it has different truth-conditions and information goes in different files.

Finally, the distinction between regular and indexed files allows us to give a simple explanation of how we manage to keep separate, in our minds, the different kind of

[229]This idea is similar to the account proposed by Thomasson (2003b).

[230]See Voltolini (2010: 108).

information we collect about a literary character. Consider this example: George, a small child, after listening to the adventures of Pinocchio that the mother told him, is convinced of the existence of a puppet named Pinocchio, who then becomes a real kid. In other words, George has opened in his mind a file PINOCCHIO and believes, mistakenly, to have acquaintance, however indirect, with the puppet of the history. But the mother corrects him and says that Pinocchio actually never existed, it is just a figment of Collodi's imagination, the author who wrote the book. At this point, what may George do with his mental files? One answer could be that George adds the predicates "not existing", "being a fictional character", "being invented by Collodi" to those already contained in PINOCCHIO. However, in the file there would be different kinds of information: some concerning the fiction in which Pinocchio is the protagonist, others, the metafictional information, that have nothing to do with the plot of the story. As we said, it seems unlikely that these pieces of information can be stored on the same level – in the same box – without generating contradictions or at least cognitive difficulties. Indeed, how could Pinocchio be made up both by Collodi and by Geppetto? One should then postulate a way to keep separate bodies of information of different nature contained in one single file. The alternative option that I want to suggest is the following: George transfers the content of the regular file into a vicarious file indexed to the story, i.e. <PINOCCHIO, *The Adventures of Pinocchio*>. Still, in his regular file PINOCCHIO, he keeps the metafictional information that now he knows about the character of Collodi's book. This explains why George can, at the same time, imagine that Pinocchio is a puppet who tells lies and yet know that Pinocchio does not exist in the real world. In the first case, George uses information encapsulated in the vicarious file, and think about Pinocchio as the puppet depicted by the novel; in the second case, he uses information encapsulated in the regular file, by thinking about Pinocchio as an abstract object.

Usually, we have both the vicarious and the regular file, linked together. But it is possible to only possess either the vicarious file alone, or the regular file alone. The first case occurs, for example, when we are not sure about the status of a character presented into a novel. For instance, in Tolstoy's *War and Peace*, literary criticism has highlighted that many characters that we used to believe as being purely fictional are in fact historical figures of the Russian aristocracy who lived for real[231]. Thus, a not perfectly informed reader may wonder whether a character like Nikolaj Bolkonskij is fictional or real. Such a reader can suspend the judgment by only keeping in her mind the vicarious file with the information drawn from the

[231] I thank Carola Barbero for this example.

narrative, i.e. about Nikolaj Bolkonskij so as depicted in *War and Peace*, without inquiring too much about the historical background. The second case occurs, for example, when a subject is aware of the existence of a novel written by Gustave Flaubert, whose main character is a certain Emma Bovary, but she has never read the book nor she knows anything about the plot. Therefore, the subject only has few metafictional information on Emma Bovary, conceived as a fictional character, that she stores in the regular file BOVARY.

A final clarification is needed. Here, we have to recall the distinction between objects of fiction and objects of error we presented in the first chapter[232]. As I said, as far as literary characters are concerned, we can take for granted that, at least in most cases, subjects possess both the two files. But when we deal with objects of error, it is much more doubtful that regular files are involved too. After all, errors are errors: they occur when one supposes an object to exist where no object does really exist. If we understand errors for what they are, they do not seem to carry ontological commitment to any sort of entities. Then, free-wheeling, or unloaded, use of indexed file appears to be the best way to account for them[233]. At least, until errors themselves become objects of some fiction, as it happened for Vulcan.

3.11 Conclusion

My work has come to an end. After having examined the current philosophical debate about fictional objects, since no theory succeeds in prevailing over the others, I suggested to put aside the metaphysical and ontological questions and tackle the issue from a cognitive perspective, by means of the conceptual tools provided by the theory of mental file. I have tried to show how much versatile and useful for our purposes is the notion of indexed file. Due to their meta-representational function, indexed files, by their own, do not need reference, but they have existential import when they are vertically linked with other files.

[232] See ch.1.1.

[233] In fact, our mental representation of an object of error necessarily pass through the filing system of some other subject, the one who makes the mistake. Otherwise, if we deployed a regular file, the object would no longer be represented as an error, since we would believe it to exist.

They enable us to store information derived from a fictional story and to keep it separate from what we know of the actual world: this information is taken to be about a real individual if the indexed file is linked to a regular one, otherwise it is not. I suggested that files gather into networks, according to the causal relations between one work of fiction and another. However, my idea of network differs from those already proposed in the philosophical literature, for instance by Perry, Everett and Friend, since networks do not involve regular files, but indeed indexed ones. By means of this conceptual apparatus, we can deal with the identity of literary characters at the fictional as well as at the metafictional level. We can address the intuition that there are as many different manifestations of a literary character as there are fictional works in which it appears: each knot of the network is provided by a file, indexed to one of the fictions in which the character in question appears. We can also address the intuition that there is a single fictional character that moves from one story to another. Indeed, since indexed files are embedded into the same network through vertical linkings, a unique referent is presupposed: all of them would have the same referent, if there was one. As I argued, we do not need any commitment to a general character, like "the Ulysses in general", collecting the essential attributes of Ulysses, as suggested, for instance, by Orilia and Eco. Only at the end of my work, I recovered the notion of regular file, in order to account for the difference, at the cognitive level, between conniving and non-conniving uses of fictional sentences. I suggested that regular files refer to fictional entities not as flesh-and-blood individuals, like indexed files do, but as abstract objects. It was not the purpose of this work to provide a theory about the metaphysical nature of fictional objects. However, now that I have an account of how people conceive literary characters in their mind, I can start from this base to develop a more precise analysis of the ontological and metaphysical aspect of the topic. This may be the point of departure for future investigations.

Acknowledgements

My sincere gratitude goes to my supervisor Carola Barbero for her continuous support, motivation and patience. Her insightful comments and guidance helped me during the time of research and writing of this thesis and indicated intriguing paths of philosophical analysis.

I am deeply indebted with two institutions. I would like to thank the *Northwestern Italian Philosophy Consortium* for giving me the opportunity to devote my time to study and research. I am grateful for all the seminars and classes that I followed during these years. A special place in my memory is occupied by the lessons on "The Philosophy of Fiction", held at the University of Torino in 2015, without which this thesis would not have been possible. I also owe much gratitude to the *Institut Jean Nicod*, where I spent many months working on my thesis. It was a truly stimulating and enriching experience.

I would like to express my gratitude to Alberto Voltolini for his encouragement and the discussions which urged me to widen my research from various perspectives. A sincere thank also goes to Carlo Penco and Tommaso Piazza.

A special acknowledge is devoted to François Recanati for all the time he spent for me, even at the cost of losing dinner and settling for a frugal meal on Rue Gay-Lussac.

I would like to thank the people I met during these years, at conferences or on other occasions, that gave me suggestions and encouragement, in particular Enrico Terrone, Manuel García-Carpintero, Stacie Friend, Zsófia Zvolenszky, Petr Koťátko.

Sincere gratitude to my colleagues: Clara Madaro, Enrico Cipriani, Marta Benenti, Emiliano Loria, Matteo Cresti, Cristina Nencha, Antonio Negro, Agostino Pinnapintor. Last but not least, I would like to thank my family for never making me miss love and support.

BIBLIOGRAPHY

_ Adams, F.; Fuller, G. & Stecker, R. (1997), "The Semantics of Fictional Names", *Pacific Philosophical Quarterly*, 78: 128–148.

_ Bach, K. (1987), *Thought and Reference*, Oxford: Oxford University Press.

_ Barbero, C. (2005), *Madame Bovary: Something like a melody*, Milano: Alboversorio.

_ Berto, F. (2011), "Modal Meinongianism and Fiction: the Best of Three Worlds", *Philosophical Studies*, 152: 313–334.

_ Boghossian, P. A. (1994), "The transparency of mental content", *Philosophical Perspectives* 8: 33-50.

_ Bonomi, A. (1994), *Lo spirito della narrazione*, Milano: Bompiani.

_ Bonomi, A. (2008), "Fictional Contexts", in P. Bouquet et al. (eds.), *Perspectives on Contexts*, Standford: CSLI Publications, 215-250.

_ Breckenridge, W. & Magidor, O. (2012), "Arbitrary reference", *Philosophical Studies* 158(3): 377-400.

_ Brock, S. (2002), "Fictionalism about Fictional Characters", *Noûs*, 36: 1–21.

_ Brock, S. (2010), "The Creationist Fiction: the Case against Creationism about Fictional Characters", *Philosophical Review*, 119: 337–64.

_ Burge, T. (1973), "Reference and proper names", *Journal of Philosophy*, 70(14): 425-39.

_ Burge, T. (1979), "Individualism and the Mental," in *Midwest Studies in Philosophy* 4: 73-121.

_ Campbell, J. (2006), "Sortals and the binding problem", in F. McBride (ed.) *Identity and modality*, Oxford: Clarendon Press, 203-218.

_ Castañeda, H-N. (1975), "Individuation and Non-Identity: A New Look", *American*

Philosophical Quarterly 12, 131–40.

_ Castañeda, H-N. (1989a), "Fiction and reality: Their fundamental connections: An essay on the ontology of total experience", in *Thinking, Language, and Experience*, Minneapolis: University of Minnesota Press, 176-205.

_ Castañeda, H-N. (1989b), "Thinking and the Structure of the World" in *Thinking, Language, and Experience*, Minneapolis: University of Minnesota Press, 235-261

_ Castañeda, H.-N. (1999), *The Phenomeno-Logic of the I: Essays on Self-Consciousness*, Bloomington, Ind.: Indiana University Press.

_ Coffa, J. A. (1993), *The semantic tradition from Kant to Carnap: To the Vienna station*, Cambridge: Cambridge University Press.

_ Coliva A. & Belleri D. (2013), "Some Observations on François Recanati's Mental Files", *Disputatio*, 36(5): 108-117.

_ Crane, T. (2013), *The Objects of Thought*, Oxford: Oxford University Press.

_ Crimmins, M. (1998), "Hesperus and Phosphorus: Sense, Pretense, and Reference", *The Philosophical Review*, 107: 1–47.

_ Cumming S. (2016) "Names", *The Stanford Encyclopedia of Philosophy* (Fall 2016 Edition), Edward N. Zalta (ed.)

URL = <https://plato.stanford.edu/archives/fall2016/entries/names/>

_ Currie, G. (1990), *The Nature of Fiction*, Cambridge: Cambridge University Press.

_ Currie, G. (1995), "Unreliability refigured: Narrative in literature and film", *The Journal of Aesthetics and Art Criticism 53*(1): 19-29.

_ Currie, G. (2010), *Narratives and narrators: A philosophy of stories*, Oxford: Oxford University Press.

_ Devitt, M. (1981), *Designation*, New York: Columbia University Press.

_ Donnellan, K. S. (1966), "Reference and definite descriptions", *The philosophical review* 75.3: 281-304. Reprinted in S. Schwartz (ed.) (1977), *Naming, Necessity, and Natural Kinds*, Ithaca, NY: Cornell University Press, 42 – 65.

_ Donnellan, K. S. (1970), "Proper Names and Identifying Descriptions", *Synthese* 21: 335-358.

_ Donnellan K. S. (1974), "Speaking of Nothing", *The Philosophical Review*, 83 (1): 3-31.

_ Donnellan K. S. (1990), "Belief and the identify of reference", in *Midwest Studies in Philosophy Propositional Attitudes*, Stanford, CA: CSLI, 201-214.

_ Eco, U. (1979), *Lector in fabula: la cooperazione interpretativa nei testi narrativi*, Milano: Bompiani.

_ Eco, U. (2000), "Dove sta Cappuccetto Rosso?", in G. Usberti (ed.), *Modi dell'oggettività. Saggi in onore di Andrea Bonomi*. Milano: Bompiani (2000): 137-157.

_ Eklund, M. (2017), "Fictionalism", *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition), Edward N. Zalta (ed.), URL =

<https://plato.stanford.edu/archives/win2017/entries/fictionalism/>

_ Evans, G. (1973), "The Causal Theory of Names", *Proceedings of the Aristotelian Society Supplementary* 47: 187-208.

_ Evans, G. (1981), "Understanding Demonstratives". Reprinted (1985) in *Collected Papers*, Oxford: Oxford University Press, 291–321.

_ Evans, G. (1982), *The Varieties of Reference*, Oxford: Clarendon Press.

_ Everett, A. (2003), "Empty Names and 'Gappy' Propositions", *Philosophical Studies* 116(1): 1-36

_ Everett, A. (2013), *The nonexistent*, Oxford: Oxford University Press.

_ Fine, K. (1982), "The Problem of Non-Existents", *Topoi* 1(1): 97-140.

_ Fine, K. (1995), "Ontological Dependence," *Proceedings of the Aristotelian Society* 95, 269–289.

_ Fine, K. (2007), *Semantic Relationism*, Oxford: Blackwell.

_ Fitch, G. & Nelson, M. (2016), "Singular Propositions", *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), Edward N. Zalta (ed.)

URL = <https://plato.stanford.edu/archives/win2016/entries/propositions-singular/>

_ Frege, G. (1892), "Über Sinn und Bedeutung", in *Zeitschrift für Philosophie und philosophische Kritik*, 100: 25–50. Translated as "Sense and Reference" in *The*

Philosophical Review, Vol. 57, No. 3 (May, 1948): 209-230.

_ Frege, G. (1897), "Logic", in H. Hermes, F. Kambartel, and F. Kaulbach (eds.) (1979), *Posthumous Writings*, Oxford: Basil Blackwell, 126–51.

_ Frege, G. (1914), "Logic in Mathematics", in H. Hermes, F. Kambartel, and F. Kaulbach (eds.) (1979) *Posthumous Writings*, Oxford: Basil Blackwell, 203–50.

_ Frege, G. (1918), "The Thought: A Logical Inquiry". English translation by A. and M. Quinton in P. Strawson (ed.) (1967) *Philosophical Logic*, Oxford: Oxford University Press, 17-38.

_ Friend, S. (2000), "Real people in unreal contexts", in A. Everett and T. Hofweber (eds.) *Empty Names, Fiction and the Puzzles of Non-existence,* Stanford (CA)*:* CSLI Publications: 183-203.

_ Friend, S. (2011), "The great beetle debate", *Philosophical Studies* 153: 183-211

_ Friend, S. (2014), "Notions of nothing", *Empty representations: Reference and non-existence:* 307-332.

_ Friend, S. (2017), "The real foundation of fictional worlds", *Australasian Journal of Philosophy 95*(1): 29-42.

_ García-Carpintero, M. (2016a), "Pretend Reference and Coreference", *Indirect Reports and Pragmatics*, Springer International Publishing: 333-358.

_ García-Carpintero, M. (2016b), "Recent Debates on Learning from Fiction", *Teorema* 35(3): 5-20.

_ Geach, P. (1969), "The Perils of Pauline", *Revue of Metaphysics*, 23(2): 287–300.

_ Geach, P. (1972), *Logic Matters*, Oxford: Blackwell.

_ Geach, P. (1967), "Intentional Identity", *Journal of Philosophy* 74 : 627-632.

_ Grice, P. (1969), "Vacuous Names", in D. Davidson and Hintikka (eds,) *Words and Objections*, Dordrecht: Reidel, 118-145.

_ Hall, K. J. (2013), "Acquaintance and Mental Files", *Disputatio* 36(5): 120-132.

_ Horsten, L. (2018), "Philosophy of Mathematics", *The Stanford Encyclopedia of Philosophy* (Spring 2018 Edition), Edward N. Zalta (ed.), forthcoming URL =

<https://plato.stanford.edu/archives/spr2018/entries/philosophy-mathematics/>.

_ Husserl, E. (1948), *Erfahrung und Urteil*, Claassen & Goverts, Hamburg; transl. by J.S. Churchill and K. Ameriks, Routledge, London 1972.

_ Ichino, A. & Currie, G. (2017), "Truth and trust in fiction", in E. Sullivan-Bissett, H. Bradley and P. Noordhof (eds.), *Art and Belief*, Oxford: Oxford University Press.

_ Ingarden, R. (1931), *Das Literarische Kunstwerk*, Tübingen: Niemeyer.

_ Ingram, D. & Tallant, J. (2018), "Presentism", *The Stanford Encyclopedia of Philosophy* (Spring 2018 Edition), Edward N. Zalta (ed.), forthcoming URL = <https://plato.stanford.edu/archives/spr2018/entries/presentism/>

_ Jackson, F. (1998), "Reference and Description Revisited", *Philosophical Perspectives* 12: 201-218.

_ Jacquette, D. (1989), "Mally's Heresy and the Logic of Meinong's Object Theory", *History and Philosophy of Logic*, 10: 1–14.

_ Jacquette, D. (1996), *Meinongian Logic: the Semantics of Existence and Nonexistence*, Berlin and New York: de Gruyter.

_ Jeshion, R. (2010), "Singular Thought: Acquaintance, Semantic Instrumentalism, and Cognitivism", in R. Jeshion (ed.) *New Essays on Singular Thought*, Oxford: Clarendon Press, 105-140.

_ Kaplan, D. (1973), "Bob and Carol and Ted and Alice", in K.J.I. Hintikka, J.M.E. Moravcsik, & P. Suppes (eds.) *Approaches to Natural Language*, Dordrecht: Reidel, 490–518.

_ Kaplan D. (1977), *Demonstratives: An Essay on the Semantics, Logic, Metaphysics, and Epistemology of Demonstratives and Other Indexicals*. Reprinted (1989) "Demonstratives", in J. Almog, J. Perry and H. Wettstein, (eds.), *Themes from Kaplan*, Oxford: Oxford University Press, 481–614.

_ Kaplan, D. (1989), "Afterthoughts", in Among et al. (eds.) *Themes from Kaplan*, Oxford: Oxford University Press, 565-614.

_ Kaplan, D. (2005), "Reading 'On Denoting' on its Centenary", *Mind*, 114: 933–1003.

_ Kripke, S. A (1979), "A puzzle about belief", in A. Margalit (ed.) *Meaning and Use*, Dordrecht: Reidel, 239-283.

_ Kripke S. A. (1980), *Naming and Necessity*, Oxford: Blackwell [originally delivered as a series of lecture at Princeton University in 1970].

_ Kripke, S. (2013), *Reference and Existence: the John Locke lectures*, Oxford: Oxford University Press.

_ Kroon, F. (1992), "Was Meinong Only Pretending?," *Philosophy and Phenomenological Research 52*, 499–527.

_ Kroon, F. (1994), "Make-Believe and Fictional Reference", *The Journal of Aesthetics and Art Criticism*, 42: 207–214.

_ Künne, W. (1982), "Criteria of Abstractness," in B. Smith (ed.), *Parts and Moments. Studies in Logic and Formal Ontology*, Munich: Philosophia Verlag, 401–37.

_ Lamarque, P. & Olsen, S.H. (1994), *Truth, Fiction, and Literature*, Oxford: Clarendon Press.

_ Landini, G. (1990), "How to Russell Another Meinongian: A Russellian Theory of Fictional Objects Versus Zalta's Theory of Abstract Objects," *Grazer Philosophische Studien* 37: 93–122.

_ Leslie, A. M. (1987), "Pretense and Representation: The Origins of Theory of Mind", *Psychological review*, 94 (4), 412-426.

_ Lewis, D. (1978), "Truth in Fiction", *American Philosophical Quarterly*, 15: 37–46. Reprinted in D. Lewis (1983), *Philosophical Papers* (Volume 1), Oxford: Oxford University Press, 261–275.

_ Lewis, D. (1983), "Postscripts to 'Truth in Fiction' ", in D. Lewis, *Philosophical Papers* (Volume 1), Oxford: Oxford University Press, 276–80.

_ Lewis, D. (1986), *On the Plurality of Worlds*, Oxford: Blackwell.

_ Lewis, D. (1999), *Papers in metaphysics and epistemology* Vol. 2, Cambridge: Cambridge University Press.

_ Lo Guercio, N. (2015), "Mental Files and Metafictive Utterances", *Kriterion* 56(132): 541-555.

_ Loar, B. (1987), "Social Content and Psychological Content", in R. H. Grimm and D. D. Merrill (eds) *Contents of Thought*, Tucson: University of Arizona Press, 99–110.

_ Mally, E. (1912), *Gegenstandtheoretische Grundlagen der Logik und Logistik*, Leipzig: Barth.

_ Marcus, R. B. (1947), "The Identity of Individuals in a Strict Functional Calculus of Second Order", *Journal of Symbolic Logic*, 12(1): 12–15.

_ Marek, J. (2013), "Alexius Meinong", *The Stanford Encyclopedia of Philosophy* (Fall 2013 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2013/entries/meinong/>

_ Mates, B. (1952), "Synonymity", in L. Linsky (ed.) *Semantics and the Philosophy of Language*, Urbana, Ill., University of Illinois Press, 111–38.

_ McKay, T. & Nelson M. (2014), "Propositional Attitude Reports", *The Stanford Encyclopedia of Philosophy* (Spring 2014 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2014/entries/prop-attitude-reports/>

_ Meini, C. & Voltolini, A. (2010), "How Pretence Can Really Be Metarepresentational", *Mind & Society*, 9 (1): 31-58.

_ Meinong, A. (1904), "Über Gegenstandtheorie", in A. Meinong (ed.), *Untersuchungen zur Gegenstandtheorie und Psychologie*, Leipzig: Barth. Reprinted in *Gesammelte Abhandlungen* (Gesamtausgabe bd. II), Graz: Akademische Druck- und Verlagsanstalt, 1971, 481-535. Transl. as "The Theory of Objects" in Roderick M. Chisholm (ed.), *Realism and the Background of Phenomenology*, Glencoe: Free Press, 1960, 76–117.

_ Meinong, A. (1910), *Über Annahmen*, revised ed. of Meinong 1902, Leipzig: J. A. Barth. 1910. Reprinted in Meinong 1968–78, Vol. IV: 1-384. Transl. with an introduction by James Heanue as *On Assumptions*, Berkeley, Los Angeles, London: University of California Press, 1983.

_ Meinong, A. (1916), *Über Möglichkeit und Wahrscheinlichkeit*, Leipzig: Barth. Reprinted in *Gesammelte Abhandlungen* (Gesamtausgabe bd. VI), Graz: Akademische Druck- und Verlagsanstalt, 1972.

_ Meinong, A. (1917), "Über emotionale Präsentation", in *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften in Wien. Philosophisch-historische Klasse*, Band [*Vol.*] 183, Abhandlung [*Article*] 2, Wien. Separate Edition: Wien: A. Hölder, 1917, 183 pp. Reprinted in Meinong 1968–78, Vol. III: 283–467. Transl. with an

introduction by M.-L. Schubert Kalsi as *On Emotional Presentation*, Evanston: Northwestern University Press, 1972.

_ Meinong, A. (1968–78), *Alexius Meinong Gesamtausgabe* [*Complete Edition*], ed. by Rudolf Haller, Rudolf Kindinger in collaboration with Roderick M. Chisholm, 7 vols., Graz: Akademische Druck- u. Verlagsanstalt.

_ Mill, J. S. (1867), *A System of Logic*, London: Longmans. Reprinted by J. Robson (ed.) (1973) *The Collected Works of J. S. Mill* (Volumes 7–8), Toronto: University of Toronto Press.

_ Motoarca, I-R. (2014), "Fictional Surrogates", *Philosophia* 42(4): 1033-1053.

_ Ninan, D. (2015), "On Recanati' s Mental Files", *Inquiry* 58(4): 368-77.

_ Orenstein, A. (2003), "Fiction, Propositional Attitudes, and Some Truths about Falsehood", *Dialectica*, 57: 177–190.

_ Orilia, F. (2002), *Ulisse, il quadrato rotondo e l'attuale re di Francia*. Pisa: Edizioni ETS.

_ Pagin, P. (2013), "The Cognitiva Significance of Mental Files, *Disputatio*, 5(36): 133-45.

_ Papineau, D. (2006), "Phenomenal and Perceptual Concepts", in T. Alter and S. Walter (ed.) *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*, New York: Oxford University Press, 111-44.

_ Parsons, T. (1980), *Nonexistent Objects*, New Haven: Yale University Press.

_ Perner, J. & Brandl, J. (2005), "File Change Semantics for Preschoolers: Alternative Naming and Belief Understanding", *Interaction Studies* 6 : 483-501.

_ Perner, J., Rendl B. & Garnham A. (2007), "Objects of Desire, Thought, and Reality: Problems of Anchoring Discourse Referents in Development", *Mind and Language* 22 : 475- 513.

_ Perner, J. & Leahy, B. (2016), "Mental files in development: dual naming, false belief, identity and intensionality", *Review of Philosophy and Psychology* 7(2): 491-508.

_ Perry, J. (1977), "Frege on Demonstratives", *Philosophical Review*, 86: 474–97.

_ Perry, J. (1979), "The Problem of the Essential Indexical", *Noûs*, 13: 3–21. Reprinted

(1993), New York: Oxford University Press.

_ Perry, J. (1980), "A Problem About Continued Belief", in *Pacific Philosophical Quarterly*, 61: 317 – 332.

_ Perry, J. (2000), *The Problem of the Essential Indexical and Other Essays*, expanded edition, Stanford, CA: CSLI Publications.

_ Perry, J. (2001a), *Reference and Reflexivity*, Stanford, CA: CSLI Publications.

_ Perry, J. (2001b), *Knowledge, Possibility and Consciousness*, Cambridge, Mass.: MIT Press.

_ Perry, J. (2002), *Identity, Personal Identity, and the Self*, Indianapolis: Hackett.

_ Phillips, J. (2000), "Two Theories of Fictional Discourse", *American Philosophical Quarterly*, 37: 107–119.

_ Pinillos, A. (2011), "Coreference and Meaning", *Philosophical Studies* 154 : 301-324.

_ Predelli, S. (1997), "Talk about fiction", *Erkenntnis* 46 (1): 69-77.

_ Predelli, S. (2002), " 'Holmes' and Holmes: A Millian Analysis of Names from Fiction", *Dialectica*, 56(3), 261-279.

_ Predelli, S. (2008), "Modal Monsters and Talk About Fiction", *Journal of Philosophical Logic*, 37: 277–97.

_ Priest, G. (2005), *Towards Non-Being: the Logic and Metaphysics of Intentionality*, Oxford: Clarendon Press.

_ Putnam, H. (1975), "The Meaning of 'Meaning' ", in K. Gunderson (ed.) *Language, Mind and Knowledge. Minnesota Studies in the Philosophy of Science* (Volume 7) Minneapolis: University of Minnesota Press,131–93.

_ Quine, W. (1948), "On What There Is", Review of Metaphysics. Reprinted in Quine 1953: 1-19.

_ Quine, W. (1953), *From a Logical Point of View*, Cambridge, Mass.: Harvard University Press; revised edition, 1980.

_ Quine, W. V. (1956), "Quantifiers and propositional attitudes", *Journal of Philosophy* 53 (5): 177-187.

_ Rapaport, W. J. (1978), "Meinongian Theories and a Russellian Paradox", *Noûs*, 12:

153–180.

_ Raspa, V. (2001), “Zeichen, 'schattenhafte' Ausdrücke und fiktionale Gegenstände: Meinongsche Überlegungen zu einer Semiotik des Fiktiven”, *Zeitschrift für Semiotik* 23(1): 57-77.

_ Recanati F. (1993), *Direct reference: From language to thought*, Oxford: Blackwell.

_ Recanati, F. (1998), “Talk about fiction”, *Lingua e Stile* 33(3): 1-12.

_ Recanati, F. (2000), *Oratio Obliqua, Oratio Recta: an Essay on Metarepresentation*, Cambridge (Mass): MIT Press.

_ Recanati F. (2006a), “Indexical Concepts and Compositionality”, in M. Garcia-Carpintero and J. Macia (eds.) *Two-Dimensional Semantics*, Oxford : Clarendon Press, 249-57.

_ Recanati, F. (2006b), “Reply to Voltolini”, in M. J. Frapolli (ed.), S*aying, Meaning, and Referring: Essays on François Recanati's Philosophy of Language*. Houndmills, Palgrave Mcmillan: 1-5.

_ Recanati F. (2010), “Singular thought: In Defence of Acquaintance”, in R. Jeshon (ed.) *New Essays on Singular Thought*, Oxford: Clarendon Press, 141-189.

_ Recanati, F. (2012), *Mental files*, Oxford: Oxford University Press.

_ Recanati F. (2013a), “Empty singular terms in the mental-file framework”, in M. García-Carpintero and G. Martin (ed.), *Thinking and Talking about Nothing*, Oxford: Oxford University Press.

_ Recanati F. (2013b), “Reference through Mental Files: Indexicals and Definite Descriptions”, in C. Penco and F. Domaneschi (eds.) *What Is Said and What Is Not*, Stanford: CSLI, 159-173.

_ Recanati, F. (2013c), “Mental Files: Replies to my Critics”, *Disputatio*, 36(5): 205-240.

_ Recanati F. (2014), “Mental Files and Identity”, *Mind, Values, and Metaphysics*, Springer International Publishing: 467-486.

_ Recanati F. (2016), *Mental files in flux*, Oxford: Oxford University Press.

_ Recanati, F. (2018), “Fictional, Metafictional, Parafictional”, *Proceedings of the*

Aristotelian Society 138(68): 1-31.

_ Reicher, M. (1995), "Zur Identität fiktiver Gegenstände: ein Kommentar zu Amie Thomasson", *Conceptus* 28: 93-116.

_ Reicher, M. (2005), "Russell, Meinong, and the Problem of Existent Nonexistents," in B. Linsky and G. Imaguire (eds.), *On Denoting: 1905–2005*, Philosophia, Munich: 167–193.

_ Reicher, M. (2016), "Nonexistent Objects", *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), Edward N. Zalta (ed.), URL =

<https://plato.stanford.edu/archives/win2016/entries/nonexistent-objects/>

_ Reimer, M. & Michaelson, E. (2017), "Reference", *The Stanford Encyclopedia of Philosophy* (Spring 2017 Edition), Edward N. Zalta (ed.)

URL = <https://plato.stanford.edu/archives/spr2017/entries/reference/>

_ Rorty, R. (1982), "Is There a Problem about Fictional Discourse?" in *Consequences of Pragmatism*, Brighton: The Harvester Press, 110–138.

_ Rosen, G. (2017), "Abstract Objects", *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition), Edward N. Zalta (ed.), URL =

<https://plato.stanford.edu/archives/win2017/entries/abstract-objects/>

_ Roseveare, N. T. (1982), *Mercury's Perihelion from Le Verrier to Einstein*, Oxford: Oxford University Press.

_ Routley, R. (1980), *Exploring Meinong's Jungle and Beyond*, Canberra: Australian National University.

_ Russell, B. (1903), *The Principles of Mathematics*, Cambridge: Cambridge University Press.

_ Russell, B. (1905), "On Denoting", *Mind*, 14: 479–93.

_ Russell, B. (1910), "Knowledge by acquaintance and knowledge by description", in *Proceedings of the Aristotelian Society*. Vol. 11: 108-28.

_ Russell, B. (1912), *The Principles of Philosophy*, London: Williams and Norgate.

_ Russell, B. (1914), *Our Knowledge of the External World as a Field for Scientific Method in Philosophy*, Allen & Uniwin LTD. Reprint (1993), New York: Routledge.

_ Ryan, M.-L. (1980), “Fiction, Non-Factuals and the Principle of Minimal Departure”, *Poetics* 9(4): 403-422.

_ Ryle, G. (1973), “Intentionality-Theory and the Nature of Thinking”, *Revue Internationale de Philosophie* 27: 255-265.

_ Sainsbury, M. (2005), *Reference without Referents*, Oxford: Oxford University Press.

_ Sainsbury, M. (2009), *Fiction and Fictionalism*, London: Routledge.

_ Sainsbury, R. (2010), “Fiction and Acceptance-Relative Truth, Belief, and Assertion”, in F. Lihoreau (ed.) *Truth in Fiction*, Munich: Ontos Verlag, 137–152.

_ Sainsbury, M. & Tye, M. (2012), *Seven Puzzles of Thought and How to Solve Them: An Originalist Theory of Concepts*, Oxford: Oxford University Press.

_ Salis, F. (2013), “Fictional Names and the Problem of Intersubjective Identification”, *Dialectica,* 67(3): 283-301.

_ Salmon, N. (1986), *Frege's Puzzle*, Cambridge, MA: MIT Press.

_ Salmon, N. (1998), “Nonexistence”, *Noûs*, 32: 277–319.

_ Santambrogio, M. (1992), *Forma e oggetto*, Milano: Il Saggiatore.

_ Schiffer, S. (1978), “The basis of reference”, *Erkenntnis* 13.1: 171-206.

_ Schiffer, S. (1996), “Language-Created Language-Independent Entities”, *Philosophical Topics*, 24: 149–166.

_ Schiffer, S. (2003), *The Things We Mean*, Oxford: Clarendon Press.

_ Schroeter, L. (2008), “Why Be an Anti-Individualist?”, *Philosophy and Phenomenological Research* 77.1: 105-141.

_ Searle, J. (1958), “Proper Names”, *Mind*, 67(266): 166–73.

_ Searle, J. R. (1979), “The Logical Status of Fictional Discourse”, in P.A. French et al. (eds.), *Contemporary Perspectives in the Philosophy of Language*, Minneapolis: University of Minneapolis Press, 233–243.

_ Smith, B. (1980), “Ingarden vs. Meinong on the Logic of Fiction,” *Philosophy and Phenomenological Research* 41, 93–105.

_ Stalnaker, S. (2008), *Our knowledge of the internal world*, Oxford: Oxford University

Press.

_ Stojanovic, I. & Fernandez, N. V. (2015), “Mental Files, Blow Up by Indexed Files”, *Inquiry*, 58(4): 393-407.

_ Strawson, P. (1959), *Individuals: an Essay on Descriptive Metaphysics*, London: Methuen.

_ Strawson, P. (1974), *Subject and Predicate in Logic and Grammar*, London: Methuen.

_ Teller, P. (2009), “Fictions, fictionalization, and truth in science”, in Mauricio Suárez (ed.) *Fictions in science: Philosophical essays on modeling and idealization*, London: Routledge, 235-247.

_ Terrone, E. (2017a), “On Fictional Characters as Types, *British Journal of Aesthetics*, 57(2): 161-176.

_ Terrone, E. (2017b), “The Band of Theseus: Social Individuals and Mental Files”, *Philosophy of the Social Sciences* 47(4-5): 287-310.

_ Terrone, E. (forthcoming) “Twofileness: Fictional Characters as Generators of Mental Files”.

_ Thein, K. (2009), “Artefacts, Non-Particulars and Model Particulars”, *Organon* 16 (4), 510-528.

_ Thomasson, A. L. (1999), *Fiction and Metaphysics*, Cambridge: Cambridge University Press.

_ Thomasson, A. L. (2003a), “Fictional Characters and Literary Practices”, *British Journal of Aesthetics*, 43: 138–157.

_ Thomasson, A. L. (2003b), “Speaking of Fictional Characters”, *Dialectica*, 57: 205–223.

_ Thomasson, A. (2009), “Fictional entities”, in Jaegwon Kim, Ernest Sosa and Gary Rosenkrantz (eds) *A Companion to Metaphysics*, Oxford: Blackwell, 10-18.

_ Thomson-Jones, K. (2016), *Current Controversies in Philosophy of Film*, New York, London: Routledge.

_ Twardowski, K. (1894), *Zur Lehre vom Inhalt und Gegenstand der Vorstellungen*, Wien, A. Hölder. Reprinted (1982), München: Philosophia Verlag.

_ Vaihinger, H. (1924), *The Philosophy of 'As if': A System of the Theoretical, Practical and Religious Fictions of Mankind*, London: Routledge and Kegan Paul. Reprinted by C. K. Ogden (ed.) (1968), New York: Barnes and Noble.

_ Van Inwagen, P. (1977), "Creatures of Fiction", *American Philosophical Quarterly*, 14: 299–308.

_ Voltolini, A. (2003), "How Fictional Works Are Related to Fictional Entities", *Dialectica*, 57: 225–238.

_ Voltolini, A. (2006), *How Ficta Follow Fiction: A Syncretistic Account of Fictional Entities*, Dordrecht: Springer.

_ Voltolini, A. (2009), "Consequences of schematism", *Phenomenology and the Cognitive Sciences*, 8(1), 135-150.

_ Voltolini, A. (2010), *Finzioni. Il far finta e i suoi oggetti*, Roma-Bari: Laterza.

_ Voltolini, A. (2013), "Probably the Charterhouse of Parma does not exist, possibly not even that Parma", *Humana Mente* 25, 235-261.

_ Voltolini, A. & Kroon, F. (2016), "Fiction", *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), Edward N. Zalta (ed.), URL =

<https://plato.stanford.edu/archives/win2016/entries/fiction/>

_ Yagisawa, T. (1988), "Beyond possible worlds", *Philosophical Studies*, 53(2): 175-204.

_ Yagisawa, T. (2001), "Against Creationism in Fiction", *Philosophical Perspectives* 15: 153–172.

_ Yagisawa, T. (2018), "Possible Objects", *The Stanford Encyclopedia of Philosophy* (Spring 2018 Edition), Edward N. Zalta (ed.), forthcoming URL =

<https://plato.stanford.edu/archives/spr2018/entries/possible-objects/>.

_ Walton, K. L. (1990), *Mimesis as Make-Believe: On the Foundations of the Representational Arts*, Cambridge, MA: Harvard University Press.

_ Walton, K. L. (1993), "Metaphor and Prop Oriented Make-Believe", *European Journal of Philosophy*, 1(1), 39-57.

_ Walton K. L. (2000), "Existence as Metaphor?" in A. Everett, T. Hofweber (eds.),

Empty Names, Fiction and the Puzzles of Non-existence, Stanford: CSLI Publications, 69–94.

_ Wartenberg, T. (2015), "Philosophy of Film", *The Stanford Encyclopedia of Philosophy* (Winter 2015 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2015/entries/film/>

_ Wetzel, L. (2008), *Types and tokens: An essay on universals*, Cambridge: MIT Press.

_ Williamson T. (2013), *Modal Logic as Metaphysics*, Oxford: Oxford University Press.

_ Wittgenstein L. (1922), *Tractatus Logico-Philosophicus* (TLP), 1922, C. K. Ogden (trans.), London: Routledge & Kegan Paul. Originally published as "Logisch-Philosophische Abhandlung", in *Annalen der Naturphilosophische*, XIV (3/4), 1921.

_ Wolterstorff, N. (1980), *Works and Worlds of Art*, Oxford: Clarendon Press.

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

_ Zalta, E. N. (1992), "On Mally's Alleged Heresy: A Reply", *History and Philosophy of Logic*, 13: 59-68.

_ Zalta, E. N. (2000), "Pretense Theory and Abstract Object Theory", in Everett & Hofweber (eds.) *Empty Names, Fiction and the Puzzles of Non-existence*, Stanford: CSLI Publications., 117–147.