TEXT + WORK
THE MENARD CASE

edited by
**TOMÁŠ KOBLÍŽEK, PETR KOŤÁTKO
& MARTIN POKORNÝ**

CONTENTS

PREFACE

The influence and reputation of *Pierre Menard, Author of the Quixote*, is easily comparable to the impact of groundbreaking theoretical texts. Numerous philosophers, aestheticians and theorists of literature, music, or visual arts have been induced by this short story by J.L. Borges to reconsider the status of the literary work of art and the status of artwork as such, to rethink the relationship between work and text (and analogously: work and the score, work and its physical bearer), and to make explicit the significance of the cultural context and the authorial intentions for the identity of a work or the nature of fictional entities. These philosophical reflections were produced by the eccentric literary aspirations of Borges' hero, as the goal of Menard's plan is to create an original literary work whose text will be, word for word, identical with the text of Cervantes' *Don Quixote*.

To construe a fully independent literary work around a purely intellectual scaffolding is no trivial task. The narrative approach that Borges chose is an appealing one. The entire case is made accessible by a narrator with his own vested interests, driven by vanity so strong that it is clear his description will never quite convey the originality, the seriousness and the significance of Menard's project. Yet all this only makes the significance of Menard's plan clearer to the reader – whereas the author is safely protected from any suspicion of naively enthusing over his own discovery or of being enchanted by his own idea. Such artifice is inaccessible to the authors of the essays included here. Fortunately, the genre of an academic paper provides them with no room for taking an ironical distance from their own arguments. In this way they have become potential targets for the sarcastic remark, once addressed to Arthur Danto, that by drawing serious consequences from Borges' little joke they demonstrate their lack of humor. However, this should not discourage anyone. No mat-

ter whether we are intentionalists or anti-intentionalists regarding literary interpretation, we can surely agree that a literary work may have unintended non-literary implications and that the value of any theoretical arguments that it provokes is independent both from the intentions of the empirical author and from the (impersonal) intention of the work, whatever way this latter intention is construed. This is why in Menardian discussions we can often see Menard's case reduced, complemented, told anew from scratch, transferred into various other art genres, generalized and otherwise emancipated from the original literary context. This occurs – in a sequentially increasing degree – in the present volume. The essays move from analyses of the identity of a literary work of art (and the problem of the relationship between text and work), as it is explicitly established by Borges' narrator, to arguments that simply employ the Menard case as an opportunity for discussing broader issues of literary studies and philosophy of literature. Select essays even abandon the field of literature altogether and move on to analogous issues in the theory of visual arts and music.

Tomáš Koblížek
Petr Koťátko
Martin Pokorný

TOMÁŠ HŘÍBEK

The Menard Case
& the Identity of a Literary Work of Art

Abstract: The essay provides a close analysis of Nelson Goodman's arguments in support of identifying a literary work with its text—the position known as *textualism*. This position results in a rejection of a widely shared notion that Pierre Menard created a new literary work of art. Opposed to textualism is *interpretationism*, a theory put forward by Arthur Danto. According to interpretationism, a literary work is an interpretation of its text, in consequence of which multiple interpretations of the same text count as different works, Menard's *Quixote* among countless others. However, Goodman and Danto no longer appear to be opposed to each other once we take into account Goodman's somewhat later suggestion that a physicalistically (i.e., syntactically) same item could perform a variety of functions. Both Danto and Goodman can then be seen as defending variants of *aesthetic functionalism*—the view that a (literary) literary work can be identified in purely functional terms. However, aesthetic functionalism falls a victim to similar objections as its better known cousin in the philosophy of mind, and the most serious of these objections come from the theory of psychological externalism offered by Tyler Burge.
Keywords: Nelson Goodman, Arthur Danto, textualism, interpretationism, functionalism, externalism.

1. What's the use of a philosophical theory of the identity of a literary work of art – the theory that seeks necessary and sufficient conditions of the identity of a literary work, according to which the works *X* and *Y* are one and the same literary work of art? To be sure, such a theory can be of use to philosophers, if we admit that philosophers have a right to ponder even those questions whose import outside of philosophy is doubtful. However, many of us think that if a philosophical theory sought answers to questions that never arose outside of philosophy, then the legitimacy of philosophy would somewhat suffer. If we wish to legitimate certain philosophical questions and theories that attempt to answer them, we must try to show their vital importance beyond philosophy. In particular, if we wish to legitimate a philosophical interest in the identity of a literary work of art, we should start by examining the disciplines such as literary

theory and criticism, in order to find out whether the questions concerning identity ever arise in them. Fortunately, for those of us to whom the practical import of philosophy matters, a cursory look into literary history and criticism suffices to demonstrate that the issue of identity lies at the very foundation of these disciplines. The students of literature have often to do with textual variations among whom they need to identify the text of the given work.[1] In many cases, it's not apparent that the work has only one, so to speak authentic, text.

In practice, we see at least two sorts of cases. On the one hand, it happens that an author produces two or more texts that more or less differ from each other. This opens a room for debate about whether each of these texts is a different variant of the same literary work, or rather a different work. For example, the Czech writer Bohumil Hrabal used to constantly rewrite his works. He had various motivations: he wished to improve his works, but he also wanted to make sure, by means of a sort of self-censorship, that his works would get published in the Communist Czechoslovakia at all. When literary historians have access to both the self-censored text as well as the original text, they usually prefer the latter. As far as I know, both of these kinds of text are usually considered to be the texts of the same work. This means that the uncensored text is usually regarded as an authentic text of Hrabal's work, yet the practice of literary history and criticism clearly admits that the work can survive more or less extensive tinkerings with the text.

---

[1] The meaning of the term 'text' should not be regarded as transparent. As we shall see later, the definition of this term is a subject of philosophical dispute. It seems, however, that common sense works with a certain notion of text which can be summed up in two points. First, 'text' refers to a sequence of words and other linguistic symbols and these symbols are understood as both syntactic and semantic units. Second, the same text can exist in an unlimited number of exemplars or tokens. In practice, we are dealing with the tokens of a text, not with the text itself, which is, ontologically, a type. 'Snow is white' and '*Snow is white*' are thus two tokens of the same text type, since they consist of the syntactically identical sequence of symbols, which also carries the same meaning; on the other hand, 'Der Schnee ist weiß' is not, despite the sameness of meaning, a token of the same text type as the previous two examples, since there is a syntactic difference. I believe that this is how the term 'text' is understood by ordinary speakers as well as literary theorists. Although it is commonly supposed that two editions of a novel share the same text despite the fact that they differ in typeface, a translation is regarded as a different text, even though it shares its typeface with the edition in the original language.

On the other hand, we can see many cases in which a text is subjected to editorial interventions beyond the control of its author. This often happens with literary classics that need to be either adjusted to the current grammatical norms or corrected for the errors left by the original editors, typesetters, or even by the author herself. This doesn't happen as often in English, but with the language that have changed more dramatically over the past century or two, such as the Czech, the texts from the nineteenth or even early twentieth centuries are not considered grammatically correct by current standards. The updating of spelling that is hardly ever limited to mere punctuation is thus a regular occupation of the editors of the older works of the Czech literature. (For example, the title of the key work of the Czech Romanticism, K.H. Mácha's poem *May* [1836], is no longer spelled *Mág*, but *Máj*.) When contemporary editors change the text of classic works, they assume, again, that the work is not identical with its text. Moreover, the 'authentic text' is not necessarily identical with a text presumably intended by its author. Quite likely, the term rather refers to the best text – the best by the current standards. And yet, contemporary historians of literature do not speak of *Mág* and *Máj* as two different works. They consider them one and the same work in two textual – more precisely, syntactic – versions.

It is important to notice, however, that once we admit such a possibility, we have already accepted a quite controversial philosophical theory. If Mág and Máj are one and the same poem despite the fact of their textual difference, what are the conditions of identity of a literary work? Moreover, let us realize that the theory of identity of literature determines its ontology. This means that if the identity of a literary work is not clearly determined by its textual identity, then work is not text. But then, what *is* a literary work? What kind of a *thing* is it? Let us not jump to conclusions, since we cannot decide in favor or against a particular theory of the identity of a literary work simply on the basis of the practice of literary history or criticism alone. First, it's far from clear that we shall be able to uncover from the practice of literary history or criticism – presumably, by means of some method of conceptual analysis – a consistent notion of the identity of a literary work. Second, even if we found out that literary historians and critics do assume a consistent notion of the identity of a literary work, it's not clear that this notion is going to be *philosophically* satisfactory – e.g., that

it will not commit us to a notion of the literary work as an onto-logically queer sort of entity. In order to be able to decide the question of whether or not the literary work of art is identical with a text – and to make a progress with the issue of the onto-logical status of the literary work – we must make a transition from conceptual analysis to constructive metaphysics.

2. In the previous section, I used several examples from the practice of literary history and criticism that seem to suggest that *the same literary work may exist in different texts*. How-ever, contemporary theories of the identity of a literary work haven't been necessarily stated in relation to the conceptual framework of the *actual* literary history or criticism. Instead, these theories have been crucially influenced by an example from an *imaginary* literary history, which suggests that *the same text could carry different literary works*. The author of this ex-ample is Jorge Luis Borges who, in his story 'Pierre Menard, Author of Quixote', offers a survey of the literary legacy of a fictional Symbolist poet Menard – an author, among other things, of a fragment of a novel whose text happens to be identical with the text of *Don Quixote* by Cervantes. Borges suggests that despite the textual identity, Menard did not produce simply a copy of a part of Cervantes' work, but instead wrote a new work. The reason for this difference is that Menard's *Quixote* has many properties lacked by Cervantes' *Quixote*, and vice versa:

> The fragmentary *Don Quixote* of Menard is more subtle than that of Cervantes. The latter indulges in a rather coarse opposition between tales of knighthood and the meager, provincial reality of his country; Menard chooses as 'reality' the land of Carmen dur-ing the century of Lepanto and Lope. (Borges 1962: 51)

In addition to these differences in content, Borges also draws attention to a contrast between an alleged superficial literalness of Cervantes' prose, on the one hand, and sophisticated allu-sions to intellectual fashions of his time that characterize Menard's work, on the other. These distinctions are supposed to be apparent when we compare relevant passages:

> It is a revelation to compare *Don Quixote* of Menard with that of Cervantes. The latter, for instance, wrote (*Don Quixote*, Part One, Chapter Nine):

> …truth, whose mother is history, who is the rival of time, depository of deeds, witness of the past, example and lesson to the present, and warning to the future.
>
> Written in the seventeenth century, written by the 'ingenious layman' Cervantes, this enumeration is a mere rhetorical eulogy of history. Menard, on the other hand, writes:
>
> . . . truth, whose mother is history, who is the rival of time, depository of deeds, witness of the past, example and lesson to the present, and warning to the future.
>
> History, *mother* of truth; the idea is astounding. Menard, a contemporary of William James, does not define history as an investigation of reality, but as its origin. Historical truth, for him, is not what took place; it is what we think took place. The final clauses – *example and lesson to the present, and warning to the future* – are shamelessly pragmatic. (Borges 1962: 52-53)

Finally, Borges emphasizes stylistic differences between Menard and Cervantes: "The archaic style of Menard – in the last analysis, a foreigner – suffers from a certain affectation. Not so that of his precursor, who handles easily the ordinary Spanish of his time" (ibid.: 53).

Borges' story is itself a literary work, not a philosophical paper. Despite that, we can distill from it a valid, if enthymematic, argument. Let's call it the *Borges Argument*:

> 1. If *X* and *Y* are one and the same literary work, they share all their properties.
> 2. *X* and *Y* do not share all their properties. (Suppose *X* is Menard's *Quixote* and *Y* is Cervantes' *Quixote*.)
> ∴ *X* and *Y* are two different works. (That is, Menard's *Quixote* and Cervantes' *Quixote* are two different works, two different novels.)

The validity of this simple argument is evident (formally speaking, it has the form of *modus tollens*). Thus, whoever wishes to question the conclusion of the Borges Argument must undermine the truth of at least one of its premises. (To be sure, even if the critics of the Borges Argument succeeded in showing that its conclusion does not follow from its premises, it would not necessarily establish the falsity of its conclusion. However, the conclusion would then need a different argument in its support.)

3. An ingenious simplicity of the Borges Argument probably explains its popularity among those philosophers who have during the last three decades or so worked on the philosophy of literature. While a conceptual analysis of the language of literary history or criticism does not promise a conclusive answer to the question of the identity of a literary work, the implication of the Borges Argument are quite clear. If it could be proved that both of its premises are, as a matter of fact, true, its conclusion would have to be accepted as a matter of logic. If one and the same text could really carry two different works of literature, it would logically follow that the textual identity is at least insufficient in determining the identity of a literary work. The Borges Argument thus promises, among other things, an extraordinary clarity about what needs to be done – namely, a proof or refutation of the two premises.

It is remarkable that the development of a philosophical debate about the concept of a literary work of art over the past three decades or so can really be seen as a commentary on the Borges Argument. Two positions have crystallized during this debate. On the one hand, there are those who believe that the Argument is not only valid, but also sound. Its first premise is nothing but an application of Leibniz' Principle of Identity. The second premise is also true, since, as explained by Borges, Menard's Quixote has many properties not shared by Cervantes' *Quixote* and vice versa. It follows that, despite their textual identity, they are two different works. But what are the criteria of identity of these works, if they differ from the criteria of textual identity? According to Arthur Danto, the influential defender of this position, two textually identical literary works can be distinguished precisely by reference to those semantic and stylistic properties listed by Borges when he differs Menard's work of from that of Cervantes. Danto writes:

> It is not just that the books are written at different times by different authors of different nationalities and literary intentions; these facts are not external ones: they serve to characterize the works(s) and of course to individuate them for all their graphic indiscernibility. That is to say, the works are in part constituted by their location in the history of literature as well as by their relationships to their authors ... you cannot isolate these factors from the work since they penetrate, so to speak, the *essence* of the work. (Danto 1981: 35-36; emphasis in the original)

To be sure, there is nothing extraordinary about the text of Cervantes' and Menard's novel according to Danto; the point applies quite generally that *no* work of literature is identical with its text, and that any text embodies a variety of works.

What are the consequences of this view of the identity of a literary work for its ontology? If we reject the identity of the work with its text, then it seems out of question to conceive of a literary work as some sort of a physical object. Clearly enough, the text itself is not, strictly speaking, a physical token itself; rather, it is an abstract type that can exist in multiple physical tokens.[2] If a novel, poem or story were identical with their respective texts, then we could at least say that even if they are, in and of themselves, abstract types, they are nevertheless embodied in a potentially unlimited number of physical objects (books, magnetic tapes, CD-Roms) or events (recitations). If the Borges Argument is valid, however, then in the situation when we are announcing, 'this is Cervantes' *Don Quixote*', while holding up the book that contains the text of this novel, we are saying something that is not, strictly speaking, true. However, it seems that Danto does not intend to accept an ontology that would make literary works of art totally detached from the physical world. In his view, literary works are *physical entities interpreted in a certain way*:

> In art, every new interpretation is a Copernican revolution, in the sense that each interpretation constitutes a new work, even if the object differently interpreted remains, as the skies, invariant under transformation. An object *o* is then an artwork only under an interpretation *I*, where *I* is a sort of function that transfigures *o into a work*: $I(o) = W$. Then even if *o* is a perceptual constant, variations in *I* constitute different works. (Ibid.: 125)

Danto suggests, in effect, that we can avoid an unattractive conclusion that literary works are something totally different from physical objects[3] by conceiving of the works as physical objects – as physical tokens of certain abstract textual types – of which certain *descriptions* are true. Danto often reveals that this ontology of literature draws on Wittgenstein's philosophy of

---

2 See note 1 above.

3 Or events, if the work exists, as I already suggested above, in the form of a recital—which can be the only mode of its existence in oral cultures. In the rest of this paper, however, I shall ignore the ontological sophistications of this sort and speak simply of 'objects' or 'entities'.

action. Wittgenstein asks what needs to be added to a movement of an arm in order for this movement to become a particular gesture; he answers that it suffices for the movement to figure in a particular intentional interpretation. [4] Based on Danto's favorite example of a series of perceptually indistinguishable objects,[5] we can say that among many perceptually indistinguishable arm movements one could be a greeting, another an expression of approval, and yet another a meaningless twitch, depending on whether an intentional description is available in a given circumstance, and if so, which one. Wittgenstein and Danto are thus suggesting that a change of label from

---

[4] Wittgenstein wonders: "what is left over if I subtract the fact that my arm goes up from the fact that I raise my arm?" (Wittgenstein 1958: §621). If we limit ourselves to the sphere of visual arts, it may suffice to speak of *perceptually* indiscernible objects, some of which are artworks. In philosophy of action, where we could encounter an objection that we would find out a difference between a gesture and a mere movement if we looked, so to speak, under the agent's skin, we could posit a condition of *physiological* indiscernibility. See also the next note.

[5] It is useful to note that Danto's philosophy of art deals almost exclusively with examples from the visual arts. Literature, music and other art forms are mentioned only occasionally. Danto's fundamental question – What makes one of a pair of perceptually indiscernible objects an artwork when the other isn't? – arose from his experience with examples of advanced art of the 1960s, such as pop art a minimalism that erased the traditional distinction between artworks and "ordinary things". In keeping with his fundamental question, Danto understands Borges' example of a two different literary works sharing a single text also a challenge to find properties other than perceptual—more exactly, other than syntactic – as identifying properties of a literary work. To be sure, Borges' example differs from Danto's notorious example, Warhol's *Brillo Boxes*, to the extent that in Borges' story we have two syntactically identical texts, both of whom embody literary works, while in Warhol's case we have two identical boxes for brillo pads only one of which, though, is a work of art. This difference between Borges and Danto isn't too important, though, since Danto himself uses an example of an imaginary series of red canvases, some of which are artworks – artworks profoundly different from each other, as suggested by their titles *Kierkegaard's Mood*, *Red Square*, *Nirvana* (cf. Danto 1981: 1 *et passim*). This example is analogous to Borges' case in that perceptually (or syntactically) indiscernible objects embody different works of art. Nevertheless, this parallelism between visual arts and literature should not blind us to the fact that these two art forms are importantly different in a way not emphasized by Danto. However precisely we eventually determine the ontological category, literary works are undoubtedly multiply realizable types. Any novel or story can exist in unlimited number of copies, each of whom counts as the *original* of a given work. By contrast, many sorts of visual arts – more precisely, paintings and carved sculptures, as different from prints and cast sculptures – exist in the form of concrete physical particulars (instantiations); that is why unauthorized replicas of, say, paintings, are *forgeries*.

'physiology' to that of 'intentional agency' is not governed by noticing any perceptual differences between things that are being so categorized, but merely by a difference in interpretation. Danto's ontology of art, based on a philosophy of action, can thus be called *interpretationism*.

Let us now turn to the other theory of the identity of a literary work, which also has, as we shall see, certain ontological consequences. The defenders of this alternative view do not necessarily question the validity of the Borges Argument, but rather its soundness; indeed, they consider at least one of its premises to be false. Nelson Goodman, who is a major proponent of this approach, agrees with Danto that the first premise of the Borges Argument is uncontroversial. However, Goodman challenges the second premise, i.e. the claim that Menard's *Quixote* and Cervantes' *Quixote* differ in many respects. In Goodman's view, the identity of a literary work coincides with the identity of the work's text. From this assumption and from the fact that Menard's Quixote and Cervantes' Quixote are two instances of the same text it logically follows that the two works share *all* their properties. And from this follows, finally, that despite Borges' opinion, Menard's *Quixote* and Cervantes' *Quixote* are one and the same novel, one and the same literary work. Menard is not an author of a new work, but rather of an inscription of an already existing work (though he produced this inscription without copying any of the available copies of Cervantes' novel, which is certainly an admirable accomplishment). Semantic, expressive and stylistic properties of Menard's alleged novel, pointed out by Borges, do not constitute a new work, but rather a new *interpretation* of the already existing novel by Cervantes (see Goodman – Elgin 1986: 62). Goodman is thus also suggesting that the question about the identity of a literary work has only two possible answers: either we identify a work with its text, or we identify a work with its correct interpretation. Goodman chooses the first option.[6]

---

[6] Goodman offered his theory of the identity of a literary work even before he met with the Borges Argument. In his groundbreaking study *Languages of Art* he stated the identity conditions of a literary work as follows: "Let us suppose that there are handwritten copies and many editions of a given literary work. Differences between them in style and size of script and type, in color of ink, in kind of paper, in number and layout of pages, in condition, etc., do not matter. All that matters is what may be called *sameness of spelling*: exact correspondence as sequences of letters, spaces, and punctuation marks" (Goodman 1976: 116). Although Goodman's conception of the identity of a

In order to be able to appreciate advantages or disadvantages of Goodman's choice, we need to first examine his view of textual identity. Goodman claims that textual identity "is exclusively a matter of syntax – i.e., of admissible configurations of letters, spaces and punctuation marks – quite independently of what the text means or otherwise refers to" (Goodman – Elgin 1986: 58). Thus, two inscriptions of the English word 'club' in two different English sentences are two inscriptions of the same text, despite its different meanings: in one sentence, it refers to an association of people united by a common purpose; in another, it denotes a stick for hitting balls. On the other hand, two inscriptions of the word 'chat', one in an English sentence, the other in a French one, are each of them an inscription of a different word. Not because it means something else in the two languages – an informal talk in English; a feline creature in French – but for purely syntactic reasons: first, 'chat' is pronounced differently in the two languages, and pronunciation is a syntactic feature, according to Goodman; second, the word combines with different words in English than it does in French. Against a common-sense custom – recognizable in Danto's conception of textual identity, and apparently reflected in the practice of literary criticism as well[7] – according to which the identity conditions of a text are both syntactic and semantic, Goodman offers a revisionist, purely syntactic notion of textual identity.

Let us accept, for the sake of an argument, Goodman's syntactic identification of a text. Why accept, however, his syntactic identification of a *literary work*? In a paper co-authored with Catherine Elgin, Goodman offers an argument in favor of this thesis, which can be summarized roughly as follows:[8]

1. There are only two ways of identifying a literary work: either the work is its text, or the work is its correct interpretation.
2. However we identify the work, all the interpretations are of a single text.

literary work precedes his confrontation with the Borges Argument, it is evident from the text of *Languages of Art* still hadn't contrasted his own conception with the Borgesian notion of the work as an interpretation. I don't know whether Goodman came to know about this alternative theory through Danto's discussion of Borges, or through a direct study of the Borges Argument, since Goodman nowhere in his discussion of Borges mentions Danto's name.

[7] See note 1 above.
[8] See Goodman – Elgin 1986: 53-57.

3. The text is identifiable syntactically, independently of its interpretations.

4. Each text admits many correct interpretations. This can be proved by a simple *modus tollens*:

(a) If each text has only one correct interpretation, then this interpretation is determined by the intentions of the author.

(b) Yet "even where an author's intentions are to some extent discoverable, they do not determine correctness of interpretation; for the significance of a work often diverges often diverges from, and may transcend or fall short of, what the author had in mind. [...] Whether or not the author's intention yields an interpretation, it certainly does not yield *the* interpretation of a text" (Goodman and Elgin 1986: 55; italics in the original).

5. If we identify the work with a correct interpretation of a text, then in consequence of many correct interpretations we end with as many works as there are correct interpretations.

6. This result contradicts three assumptions of an interpretive practice: (i) Correct interpretations are interpretations of the same text. (ii) Correct interpretations are interpretations of the same work. (iii) The richness of meaning, characteristic of the great works of literature, follows from the fact that the great works have many correct interpretations.

7. Since there are only two alternatives and the alternative recommending the identification of the work with a correct interpretation is inadmissible, we are left with the other alternative.

∴Therefore, the work is what remains constant throughout the process of interpretation, which is the text defined syntactically.

It's clear that a number of premises in this argument could be challenged – whether it is the assumption expressed in the first premise to the effect that there are only two ways in identifying a text; or the syntactic conception of the text, assumed in the third premise; or, finally, an appeal to the features of ordinary interpretive practice in the sixth premise. We shall have a room from a critique of these assumptions in the remaining sections of this essay. In the remainder of the present section, however, I should like to draw attention to a deeper motivation behind Goodman's notion of the identity of a literary work of art in his general ontology and philosophy of language, which are characterized by two doctrines: extensionalism and nominalism.

Extensionalism is the thesis, according to which it is possible to decide the question about the sameness of meaning without recourse to such terms as 'intension', 'proposition' and even 'meaning', provided we meant by this expression something

else than extension. [9] These expressions evoke queer entities that lack clear identity conditions. The motivation for a syntactic notion of the text should now be evident. If it's generally forbidden to posit intensional entities, then it's self-understood that we need to get by without them not only when we are concerned with the issue of the identity or difference of *meaning* of two expressions or texts, but also when we are dealing with the question of the identity or difference of an expression or text *per se*. Two expressions or texts mean the same, if they share extension; two expressions (or texts) *X* and *Y* are the same, if they are two instances of the same type of a chain of characters.

When it comes to *nominalism*, it's well known that the traditional version of this doctrine amounts to a rejection of the notion that general terms are genuine names. Only the names of concrete particulars are considered to be referential. In contrast to this traditional, Scholastic nominalism, Goodman is not afraid to admit into his ontology abstract entities. His nominalism "requires only that whatever is admitted as an entity at all be construed as an individual" (Goodman 1956: 198-199). [10] What does it mean to 'construe' something as a particular? According to Goodman, it means to reject "the composition of more then one entity out of the same entities by means of any chains of membership" (Goodman 1956: 159). Menard's *Quixote* is composed of the same syntactic units as Cervantes' *Quixote*; nominalism thus mandates regarding them as a single work. Similarly to Danto's theory of the identity of a literary work, Goodman's theory also has consequences for ontology, i.e., for the question

---

[9] Goodman defines extensionalism as a doctrine, according to which "two predicates have the same meaning if and only if they apply to the same things – in other words, have the same extension" (Goodman 1949: 5).

[10] In 'A World of Individuals' (1956), Goodman offers a brief defense of his nominalism against several objections. For an attempt at complete nominalist ontology, see his *The Structure of Appearance* (1951). It might also be helpful to take note of similarities and differences between Goodman's philosophy and the philosophy of his Harvard colleague, W.V.O. Quine. Both Goodman and Quine started off as nominalists and extensionalists. One result of their common interest in nominalist systems was a paper 'Steps Toward a Constructive Nominalism' (1948); the best-known statement of Quine's extensionalism is his 'Two Dogmas of Empiricism' (1951). While both philosophers retained their commitment to extensionalism for the rest of their respective careers, Quine eventually admitted into his ontology classes (see Quine's *Word and Object*, §55). See Goodman's 'Worlds of Individuals' (1977) for a retrospective comparison of his own and Quine's nominalisms.

'What sort of a thing is the literary work?' If we didn't take into account the special character of his nominalism, we might think that Goodman is forced to take a literary work as a concrete particular – let's say *this* copy of *Don Quixote*. Such a theory would surely be highly counterintuitive; it's clear that if I burn *this* copy of *Don Quixote*, this does not annihilate the work. Goodman is able to avoid such a conclusion, since his nominalism is not, as we have just seen, inimical to abstract entities. The only requirement is that abstract entities that will be admitted into nominalist ontology be conceivable as particulars. From this point of view we could perhaps make sense of some of Goodman's difficult claims, such as this one from *Languages of Art*: "A literary work, then, is not the compliance-class of a text but the text or script itself. All and only inscriptions and utterances of the text are instances of the work" (Goodman 1976: 209). A distinction between instances of a text and the text itself is nothing but the distinction between tokens and types, while types are clearly at a different ontological level from tokens. It seems, then, that Goodman here admits the existence of something like a text of *Don Quixote* so to speak *per se*, and he identifies the work with this entity rather than with the individual copies or recitations of the text. If this conception is compatible with nominalism, it must be possible to show that by admitting types of a text we are not cluttering the nominalist universe with entities that consist of exactly the same concatenation of elements as some other entities. Whatever shape will the nominalization of texts as abstract entities take at the end of the day, it seems clear that Goodman identifies literary works with their texts. Thus, let's call his theory *textualism*.[11]

4. In this and the following sections, I should like to critique both Danto's and Goodman's respective philosophies of literature. Rather than agreeing with one against another of these theories, I wish to point out a certain assumption that both of them share. My plan for the rest of this paper is as follows. I shall first introduce a case *à la* Borges, but with a twist that should make more probable Danto's conclusion (though not necessarily his ontology). Next, I shall briefly critique Goodman's argument leading to the conclusion that Menard's *Quixote* and

---

[11] The term 'textualism' as a label for Goodman's ontology of a literary work is coined by Gregory Currie in Currie 1991.

Cervantes' *Quixote* are one and the same work. In the conclusion of the present section, however, I shall turn to a different version of Goodman's theory that he left somewhat undeveloped, even though it suggests that he might admit that Cervantes and Menard are authors of two different literary works after all. This alternative theory of identity is part of Goodman's later general philosophy of art which, as it turns out, shares certain features with Danto's philosophy. I shall critically examine these features in the final section of this essay.

After finishing Borges' short story, we need not be convinced that Menard really authored a new literary work. After all, Borges himself is telling us in the voice of the narrator – Menard's friend and an executor of his literary legacy – that Menard once read Cervantes' novel and that he "did not want to compose another *Don Quixote* – which would be easy – but *the Don Quixote*" (Borges 1941: 48). A skeptic could point out that if Menard remembered something from his early reading of Cervantes' novel, and if Menard's ambition was to write *Don Quixote* – not some other novel, by *Don Quixote* – then these facts make Borges' contention concerning Menard's originality much less convincing.[12] However, as long as we are interested in deriving a philosophical content from Borges' idea, we need not be bothered by ambiguities of his particular fictionalization of this idea, and adjust his story in such a way that the desired conclusion follows the premises in the most convincing way.

Let's consider the following case.[13] In his novel *The Republic of Whores*, which is a satire of the life under the Communist rule in the 1950s Czechoslovakia, the Czech author Josef Škvorecký incorporated selections from the authentic propagandist poetry that had been produced by the Communist writers at the time, although in the context of Škvorecký's novel these turgid rhymes read like a literary parody.[14] But we can imagine an alternative scenario. Suppose that Škvorecký authored not

---

[12] Or much less clear, as far as Borges' own intentions are concerned. B. R. Tilghman thinks that Borges' story is an absurd prank because it describes an impossible object and that Danto demonstrates a remarkable lack of humor, if he is prepared to take Borges seriously and to mine his joke for an ontology of literature. However, Tilghman's critique is vitiated by his a priori mistrust of philosophical theory, derived from the later Wittgenstein. See Tilghman 1982.

[13] I am using a variant of the case introduced in Currie 1991: 328.

[14] Thanks to the literary historian Daniel Vojtěch for bringing this example to my attention. When I was reading Škvorecký's novel for the first time, I simply assumed that he wrote the alleged examples of the propaganda poetry himself.

only the novel, but also the propagandist poetry in it. Suppose, further, that the rhymes sharing the texts with Škvorecký's parodies were also in fact authored by a real-life Communist poet during the 1950s. (Let us say that Škvorecký could not know those Communist poems because their 1950s author had sent them to some literary magazine, forgot about them shortly afterwards, and his works got buried under some dusty papers until recently – many years after their syntactic doubles appeared at the pages of Škvorecký's satirical novel.) What lesson should we draw from such a case? It seems evident that the verses at the pages of *The Republic of Whores* are not the same verses as those that were, according to our fictional scenario, discovered under the heaps of dust in the editor's office of some long-defunct Communist literary magazine. These two sets of verses differ in terms of their literary properties. For example, they belong to two very different genres – Škvorecký's poetic creation is a parody, while its 1950s counterpart is a sincere contribution to the propagandist poetry. This alternative story supports the same conclusion as Borges' original fiction, but without obscurities and ambiguities that could be exploited by a skeptic. This variant of the Borges' case thus makes more plausible Danto's thesis that two perceptually (or syntactically, or physically) identical objects could embody two different works.

Yet Goodman and Elgin are uncompromising in the face of examples such as these. They write:

> What Menard wrote is simply another inscription of the text. Any of us can do the same, as can printing presses or photocopiers. Indeed, we are told, if infinitely many monkeys were to type for an indefinitely long time, one would eventually produce a replica of the text. That replica, we maintain, would be as much an instance of the work *Don Quixote* as Cervantes' manuscript, Menard's manuscript, and each copy of the book that ever has been or will be printed. (Goodman – Elgin 1986: 62)

We can suppose that Goodman and Elgin would reject the variant of the Borges case that I offered in the previous paragraph, since their ontological purism does not allow them to admit an ontology of possible worlds. Therefore, let us limit ourselves to their own example with the typing monkeys. Let us admit, then, that a lucky monkey one day produces a sequence of symbols syntactically identical with the text of *Don Quixote*, and that

this is going to happen in the actual world. I claim that, first, such a sequence of symbols is not an instantiation of the novel *Don Quixote*, and that, second, it is not a replica or copy of the same textual type whose instantiations are both Cervantes' and Menard's *Don Quixote*.

I shall begin with the first point. The scenario suggested by Goodman and Elgin hardly comes to pass sooner than after millions years of random attempts of countless generations of monkeys, or the evolutionary descendants of the present-day monkeys. Be it as it may, we can hardly expect that there will be in that distant future any human interpreters left to interpret the text as the novel *Don Quixote*. What is more important, there will no longer be a culture or tradition which has a room for the concept of the novel or a literary work in general. It seems clear that the concept of a literary work differs from the concept of (say) star or stone in that the former, unlike the latter, signifies a kind of entity, which needs for its existence the context of a specific culture or *audience*. The community of monkeys does not constitute such an audience; however, the latter could not be constituted by a community of intelligent Martians, who might succeed in deciphering the *linguistic* meaning of the text typed up by monkeys. Finally, if Goodman and Elgin now wished to claim that the text typed up by monkeys *could* be interpreted as *Don Quixote*, *if* it could be accessed by a member of our or a not-too-distant culture, then they would be relying on the modal vocabulary forbidden by their ontological purism. An accidentally emerged sequence of syntactically identifiable units for which there are no available interpreters cannot be regarded as an instance of an existing literary work.[15]

But let us turn to the second point – that is, let us ask whether we can consider an accidentally emerged sequence of symbols to be a token of the same *text* type, whose other to-

---

[15] Petr Koťátko reminds me that philosophical thought experiments must respect only physical laws and logical consistency, so that nothing prevents us from imagining that a monkey types up the text syntactically identical with that of *Don Quixote* already tomorrow, instead of ten million years into the future. However, I hold that the resultant text would not be a copy of the text of Cervantes's novel, or even a token of Cervantes' novel itself, because it would originate in a causal isolation from the genuine copies of Cervantes's novel, and in isolation from the novel's audience. Should Goodman wish to object that any one of us had an opportunity to interpret the monkey's text tomorrow, he would have to admit that suggestively shaped cracks in a dry ground also constitute a text.

kens are the manuscripts and copies of Cervantes' *Quixote* and Menard's *Quixote*. In other words, we are asking whether Goodman's conception of textual identity, implied by the third premise of the above-cited argument, is admissible. Suppose in that incredibly distant future when a chimp accidentally types up a sequence of symbols that are syntactically identical with a fragment of either Cervantes' or Menard's novel there will still exist some copies of the text of either of these two novels. (If not in the form of pages covered with print, then in some other, more durable medium.) I claim that in such a case, the extant copies are indeed the copies of the text of both Cervantes' and Menard's novel, but the sequence of symbols typed up by the lucky chimp is not a copy of either. It is evident that for a sequence of symbols to be a copy of a given text, two conditions must be fulfilled, and syntactic identity is only one of them; in addition to that, a sequence of symbols can be a copy of a text only if the former is *causally* related to the latter. Without the causal connection, a sequence of symbols produced by a chimp isn't a copy of the text of *Don Quixote* any more than a canvas, randomly stained by paint shot from a centrifuge, is a copy of Rembrandt's *Polish Rider*, even in the case that the two pictures were 'syntactically' indiscernible.[16]

Before concluding my analysis of the criticism of the Borges thesis that we find in Goodman and Elgin's work, I wish to ponder the sixth premise of their argument reconstructed in the previous section. Goodman and Elgin appeal to certain features of common-sense literary interpretation that is, in their opinion, recoverable from both ordinary reading and expert discourse on literature. I should like to make two comments on this. First, it's not even clear whether Goodman and Elgin got the character of a literary interpretation right. It seems that at least some of the influential schools of contemporary literary theory do not assign a unitary work to the same text.[17] Second, it looks as if Goodman and Elgin are willing to defend their revisionist syntactic conception of textual identity by appeal to a certain aspect of an ordinary notion of interpretation, namely an assumption, widely shared among ordinary readers, that various interpretations are related to the same text. But then Goodman and Elgin respect

---

[16] Danto uses this example in Danto 1981: 31-32. Cf. also Putnam 1981: 1.
[17] Here I have in mind various forms of literary deconstruction that, as far as I know, question the concept of a unitary literary work.

one aspect of ordinary use at the price of revisionism elsewhere. Against such a strategy it might be objected that the choice among the aspects of ordinary use that ought to be saved and those that are to be jettisoned is rather arbitrary. If we reject Goodman and Elgin's assumption that their conception of literary interpretation is dictated by common sense, we can produce an alternative conception of interpretation whose goal will not be (say) to constitute the literary work itself. Thus we could, by the way, arrive at a third conception of a literary work, despite Goodman and Elgin's assertion that their dichotomy – either text or interpretation – exhausts the range of possibilities.[18]

In the conclusion of this section, I should like to point out a different view of the Borges Argument that is briefly suggested by Goodman and Elgin. In the debates about the ontology of literature, Goodman is usually classified as a defender of the identification of work and text – i.e., as a textualist – who rejects the conclusion of the Borges Argument. This position perhaps represents Goodman's official view, and this is how I have so far presented his views in this essay. However, commentators usually skip over a short polemical paper from 1978, in which Goodman toys with the idea that the same text could belong to two works after all – hence, Cervantes' Quixote and Menard's Quixote could be two novels.[19] This short paper is primarily a critique of Richard Wollheim's ontology of art. Wollheim supposes that the work of art is to a physical object the way the soul is to a body in Cartesian dualism. The work of art is thus an entity *sui generis* – an aesthetic entity? – different from physical objects with which it happens to be mistaken. Goodman rejects this ontology as well as its alternative – the identity of the work and the physical object. In his view, dualism and materialism don't exhaust the range of possibilities. He defends a "functional view" (Goodman 1978a: 142) of the work of art. The passage in which he sums up this functionalist alternative is worth quoting at its full length:

> Quite clearly a piece of bronze may be a work of art or a bludgeon; a canvas may be a Rembrandt masterpiece and a blanket. But can a physical object and a work be the same? We might

---

[18] Petr Koťátko examines the acceptability of the Goodman and Elgin conception of literary interpretation in the third section of his paper 'Text, Work, Interpretation', in this volume.

[19] See Goodman 1978a.

equally well ask, but seldom do, whether the bronze and the bludgeon are the same, or the canvas and blanket. The answer indicated seems to be that in both cases we have the same physical objects performing different functions. We need not look for an aesthetic object distinct from a physical vehicle but only distinguish aesthetic from other more practical functions. While no bludgeon is a Rodin and no blanket is a Rembrandt, the same physical object may batter or warm in some contexts and perform the symbolic functions of a work of art in others. (Goodman 1978a: 142)

From the point of view of our topic, the following addendum is crucial for this theory: "Perhaps this bears also on the Borges problem, resolving it into questions about when a given text in Spanish is the Cervantes and when it is the Menard *Don Quixote.*" (Ibid.; italics in the original). This alternative solution of Borges' problem which, in contradistinction to Goodman's official view, admits that Cervantes and Menard are authors of two different novels, is thus built on the assumption that the concept of a *literary work* is similar to those of an *engine* or a *chair*. All these concepts are functional in the sense that items that fall into their respective extensions are characterized in terms of their functions, instead of their intrinsic physical constitution. An engine is whatever it is that transforms various forms of energy into mechanical movement; a chair is whatever it is that humans can sit on without having their knees higher than their buttocks, and that at the same time provides a support for our backs, if not our arms. The same item can serve many other functions, though. An item that is defective in one role can flawlessly execute another (a broken engine might be a perfect skull cracker; a three-legged chair can still serve as a perfect barrier, if we use it to prop up a door). The work of art is, in this view, any object, of whatever physical constitution, that happens to fulfill an aesthetic function.[20] Let us repeat: that which makes a given object

---

[20] Aesthetic function is, for Goodman, a variety of symbolic function. In *Languages of Art*, Goodman distinguishes among four "symptoms of the aesthetic." In other words, these are four types of aesthetic function: (1) *Syntactic density*. Goodman means that minutest differences influence the identity of aesthetic symbols. (2) Semantic density. Minutest differences influence the reference of certain aesthetic symbols. (3) *Relative repleteness.* This is the symptom of those aesthetic symbols that have relatively many significant properties. (4) *Exemplification*. A symbol is a sample of properties that it metaphorically or literally manifests. See Goodman 1976: 252-255. In *Ways of Worldmaking*, Goodman adds the fifth symptom: (5) *Manifold and complex reference.* An

into a work of art is not any of its physical properties, or other properties that would differentiate it from other physical objects, but merely a fact that this object under certain circumstances plays an aesthetic role.

Even though Goodman does not use the idiom of physical doubles, his view implies that there could be two physically indistinguishable objects, one of which would play the role of a work of art, while the other would play a much humbler role of – to use Goodman's own example – a blanket. Reverting to Goodman's other example with the monkeys that will eventually produce a sequence of symbols syntactically identical with the text of Don Quixote: from the point of view of Goodman's alternative theory, the product of the typing monkeys is not necessarily a copy of the text of the novel, because there might not be anybody around who could use this sequence of symbols in such a way. Goodman's functionalism thus meets Danto's interpretationism. Both Goodman's unofficial position with respect to the Borges' thesis and Danto's position can thus be regarded as two variants of *aesthetic functionalism*. As we saw in the third section, Danto holds that a movement isn't turned into a gesture by any sort of physical property that the given movement had to have as if in addition to its other physical properties. Similarly, a canvas covered with paint is not turned into a picture that the picture has but the colored canvas doesn't. It is possible that out of two physiologically indistinguishable movements one is a greeting, while the other a nervous twitch; and that out of two physically identical canvases one is *Kierkegaard's Mood*, while the other a mere blanket.[21] Similarly, out of two syntactically indistinguishable sequences of symbols one could be the text of Cervantes' novel, while the other a mere arabesque;[22] or one could be the Cervantes' novel, while the other the Menard's novel. In this place, I wish to emphasize a common tendency of Danto's position and Goodman's unofficial

aesthetic symbol refers in many ways to other symbols. Goodman doesn't mean that all types of artworks shares all the five aesthetic symptoms, or fulfill all these functions to the same extent. For example, syntactic density is typical for painting, but not for literature. See Goodman 1978b: 68.

[21] See note 5 above.

[22] Here I'm alluding to Wittgenstein's example with two syntactically indistinguishable sequences of symbols, one of which functions in our linguistic community as an inscription of a differential equation, whereas the other plays a mere decorative role during a religious ritual in some kind of exotic tribe. Danto cites this example in Danto (1981: 4) without giving the exact reference.

theory; undoubtedly, though, there are also significant differences between the two views.[23] I believe that the fundamental common feature of both theories is the view of the work of art as an *arbitrary property* of a certain artifact – whether it is a canvas covered with paint, a bronze cast, or a sequence of printed marks. However, as I shall show in the final section, this conception is problematic. If my argument goes through, then Goodman's unofficial view is not a significant advance over his official theory, even though the former allows a distinction between two syntactically identical works. Similarly affected is Danto's theory of literature, since Danto shares with Goodman a functionalist conception of the work of art.

5. In this final section I should like to examine Danto's and Goodman's functionalism from the perspective of contemporary philosophy of mind. What matters is not that functionalism was originally proposed as a solution of the mind-body problem, but rather that this theory was most elaborated and later also criticized within the philosophy of mind. If we wish to find out how plausible is aesthetic functionalism, it might be useful to start with a review of some arguments for and against psychological functionalism. I should also note that the following remarks on aesthetic and psychological functionalism are primarily critical; this means that the reader who is looking for a fully elaborated positive alternative to the functionalist notion of identity and nature of mental states and works of art will not find it here. Nevertheless, a criticism of the functionalist theory allows us to determine certain requirements that a fully articulated alternative to functionalism – whether psychological or aesthetic – must satisfy. One of these requirements is anti-reductionism.

The basic motivation for psychological functionalism is provided by the concept of the multiple realizability of psychological states. As we saw earlier, the definitions of functional terms do not mention the material constitution of the entities that are the

---

[23] For example, Danto and Goodman part ways on the issue which factors play a role in the identification of a certain object as a work of art. Whereas Danto emphasizes the necessity of situating the given object in the context of theory and history of art, Goodman believes that sensitiveness to his five symptoms of the aesthetic (see note 20 above) is enough. For Danto's critique of Goodman's view, see Danto 1981, chap. 2.

referents of these terms.[24] An engine and a chair can be constructed out of metal, plastic or ceramic materials. Philosophers who study the mind-body problem realized that the definitions of mental states should not mention the material constituents of the psychological subjects, either. Only a chauvinist would insist that psychological terms are applicable exclusively to the creatures made of carbon. If one day we meet aliens who show the evidence of intelligent behavior, we should describe their behavior using the usual intentional idioms, despite the fact that chemically speaking these beings will be titan based.[25] While in the case of humans the thought that p is realized (say) by means of a c-fiber firing, in the case of titan-based aliens the same type of thought is realized by means of an entirely physical process. However, the success of psychological functionalism was due not so much to its making room for hypothetical intelligent aliens, but rather to its making sense of real intelligent machines. It was their awareness that computational processes can be realized in a variety of physical systems that led philosophers to the idea that the concept of a mental state should be independent of its physical realization. Reflection on the theory and practice of information technologies led to the view that psychological terms are *formal* in nature, i.e. that they refer to formal structures rather then to material mechanisms.

From this it follows that the psychological functionalist identifies a certain type of mental state – e.g., the thought that p – from the point of view of the function or role that the state plays within a given psychological subject. It is quite likely that the general function of psychological states is to increase the survival chances of the subject. From this point of view it should be possible to define each specific type of mental state in terms of its specific contribution to the survival of the organism. For simplicity's sake we could regard the thought that p as a formal mechanism activated by a certain irritation of nerve endings and eventually culminating in the production of a certain type of behavior. Thus, the functional definition would not mention facts of physical constitution of the given subject, but it

---

[24] The following summary of psychological functionalism is based on the classic formulation of the theory in Putnam (1967).

[25] Functionalism was formulated both for qualitative mental states characterized by a certain kind of phenomenology, and for intentional mental states characterized by a propositional content. I am limiting my discussion to parallels between a functionalist notion of literary works and intentional mental states.

would rather determine the mental state as a *causal intermediary* between a certain type of input and a certain type of output. The fact that functional definitions of mental states do not mention their material constitution does not mean, of course, that mental states are something else than the states of the body of the subject; likewise, the fact that 'chair' is a functional term does not imply that the chair is something else than a material object. Functional description is a description of physical states, events or objects at a higher level of abstraction.[26]

Psychological functionalism thus promised to satisfy the desiderata of materialist ontology without its traditional chauvinism. And yet it seems that psychological functionalism must be rejected due to a devastating criticism from the point of view of a new theory of the identity of mental states, so-called psychological externalism. According to a traditional view, two subjects share a mental state – i.e., each subject realizes a token of the same type of mental state – as long as these two subjects are identical in all their internal parameters. A materialist version of this view adds that two subjects share a mental state as long as they are identical in all their internal *material* parameters. Functionalism weakens this requirement in that two subjects are mentally identical as long as their internal formal structure satisfies the same *functional description*, whether or not they coincide in material constitution.

---

[26] For completeness' sake, it's worth remembering that functionalism was offered in two major varieties. On the one hand, the so-called 'machine functionalism' construes mental states as the internal states of Turing machines. The attractiveness of Turing machines consists in the fact that they admit countless physical realizations. Thus, a certain mental state M is defined, according to this theory, only implicitly as the state of a Turing machine with respect to specific inputs and outputs. A certain Turing machine T is, in relation to specific inputs and outputs, a machine description of the psychological subject P, if and only if P realizes T. On the other hand, the so-called 'causal-theoretic functionalism' construes the names of mental states as abridged descriptions that characterize mental states from the point of view of their typical causes and effects. The network of these causes and effects can include other mental states as well. For example, the term 'pain' can be translated as 'the state usually caused by an injury to the body tissue leading to cries and states of self-pity'. A complete psychology of the subject P would generate complete functional definitions of all her psychological terms. Such a theory would thus be analogical to the Turing model of the psychology of the subject P. Machine functionalism can even be considered special case of the causal-theoretic functionalism. Machine functionalism was first proposed by Hilary Putnam in Putnam (1967); causal-theoretic functionalism is best known from the work of David Lewis, see esp. Lewis (1966) and Lewis (1972).

Externalism challenges the notion of the identity of mental states shared by all these theories by two arguments.[27] Let us imagine that the subject *X* thinks that aluminum is a light and soft metal used for the manufacturing of pots and pans, while the subject *Y* believes that arthritis affected his thighs. *X* is capable of identifying the samples of aluminum even though she is ignorant of the chemical structure of this metal, whereas *Y* knows how to recognize at least some cases of arthritis despite the fact that he errs when he thinks that this disease can affect thighs as well as ankles. Now imagine that *X*, or her physical replica, finds herself in a different environment, which differs from the actual one only in one tiny aspect: it contains a metal, which otherwise looks exactly like aluminum, except that it has a different chemical structure. Similarly, *Y*, or his exact physical twin, is in an environment which diverges from the actual one only to the extent that nobody there has yet identified arthritis as a special sort of disease (or syndrome); we can say that the extension of the term 'arthritis' in this alternative environment includes even rheumatic inflammation of muscles. In that case, X lacks aluminum thoughts in the aluminum-free environment; likewise, Y lacks arthritis thoughts in the environment with no concept of arthritis. From these two arguments there follows an alternative thesis concerning the identity conditions of psychological states. According to this thesis, for two subjects to share the same mental state it does not suffice that they are identical with respect to their relevant internal parameters; the two subjects also need to have the same relations to their respective environments – hence the label 'externalism' for this alternative theory of the identity conditions for intentional mental states. X and her physical twin do not share, despite their identity in all relevant internal respects, the same mental state, because *X* in her physical environment has causal relations to the bits of aluminum, while twin-*X* in his alternative physical environment has causal relations to the bits of a different metal. Similarly, *Y* and his twin do not share the same mental type, because the term 'arthritis' expresses in *Y*'s linguistic environment a different concept than that expressed by a syntactically and phonetically identical term in twin *Y*'s environment.

Externalism undermines psychological functionalism in the sense that two subjects can satisfy the same functional descrip-

[27] My summary of the two arguments is based on Burge (1986).

tion and realize the same functional architecture, and yet differ in the mental type: X's thought is about aluminum, while twin-X's thought is about another kind of mental; Y's thought is about arthritis, whereas twin-Y's thought is about a different disease entirely. Externalist arguments support the idea that the identity of an intentional mental state is independent of physical constitution or functional architecture. What, then, does the identity of an intentional mental state depend on? Under what conditions does a particular thought persist or change? We can notice that common sense assumes that the identity of content is decisive for the identity of an intentional mental state. A thought is the same thought as long as it has the same content; a thought ceases to be the same thought as soon as its content changes. [28] Functionalism is a revisionary theory of mind because it undermines the assumption of common sense and instead suggests that a thought remains the same as long as the functional identity of the subjects is preserved (or, as in the case of twins, X is functionally identical with her twin, and Y with his).

In order to realize the implausibility of functionalist theory of the identity of mental types, consider the following argument. [29] Functionalists assume that the same mental type can be realized in a variety of physical structures. Let us say that a mental state of a particular type will be realized by means of a c-fiber firing in a system P, while the same type of state will be tokened by an entirely different physical state or event in a system Q. Evidently, this view implies that psychological expressions are accidental or – to use an established technical term – non-rigid. Thus, according to this view, psychological expressions denote different states or events in different contexts. Yet, given the currently accepted understanding of a posteriori necessity, the functionalist view is untenable. If functionalists wish to pay their respects to physicalist ontology by expecting that neuroscience will one day discover the identity of psychological states as physical states of one sort or another, then they need to admit this: once neuroscience determines the physical nature of intentional psychological states in the actual world, it will thus fix the denota-

---

[28] The idea that the identity of mental content is essential for the identity of a mental state is emphasized by Burge, in particular in Burge (1979) and Burge (1993).

[29] I am following the argument offered by Joseph Owens in Owens 1986, esp. 163-165.

tions of psychological terms in every possible world in which these terms denote. Consider a parallel case: chemistry, having discovered the chemical structure of water in the actual world, determined the nature of this substance in every possible world.[30] If we are imagining a colorless, tasteless and odorless liquid which is not H_2O, then we are not imagining water, but some other liquid which only shares some superficial properties with water. Similarly, if we are imagining that a psychological state denoted by the description 'the thought that *p*' could have a content other than p, then we are imagining a quite different psychological state (which might share its phenomenological properties with *the thought that* p—provided that intentional states do have phenomenological properties). To be sure, from this doesn't follow that organisms of a different physical constitution, or even inorganic systems (e.g., intelligent machines) couldn't token psychological states of the same type. Rather, it follows that it's mistaken to consider psychological expressions as non-rigid designators of physical states. In other words, psychological expressions cannot be regarded as descriptions that could denote the same physical type. Such a conception of psychological expressions would be *reductionist*. Psychological expressions designate rigidly, and what they designate across different possible worlds is psychological states identified in terms of their contents.

What is the connection between this critique of psychological functionalism and the main topic of this essay, which is the identity and ontology of a literary work? In the previous sections, we have seen that, for Danto, the work of art is a physical object interpreted in a particular way; for Goodman, the work is a physical object used in a particular way. Both these theories are variants of aesthetic functionalism, since they identify the work of art in terms of a certain role, which the object in question plays, rather than its physical substrate. Two physically diverse objects can play an aesthetic role of the same type. It seems, however, that this theory runs into similar difficulties as psychological functionalism. In the aforementioned externalist thought experiments we have been considering situations, in which a subject object is transferred from one context or possible world to another, or in which each of two physical or functional twins

---

[30] For 'rigid' vs. 'non-rigid' designation and the conception of a posteriori necessities, see Kripke 1980.

is situated in a physically different situation or possible world. Similarly, we can consider a transfer of an object from one environment to another, or physical twins each situated in a somewhat different environment. As in the thought experiment, in which a subject actually misapplied the term "arthritis," while in a counterfactual situation he applied it correctly, since the two situations different in terms of a linguistic convention instead of a physical fact, a physical object could be transferred from one environment to another that differ from each other in terms of a cultural practice, rather than physically. We know that Goodman rejects an ontology of possible worlds. Yet we have seen that we can interpret his example with typing monkeys in such a way that we obtain alternative futures of the actual world, in which the identical syntactic structure, which would play the role of a text of Don Quixote, cannot play this role because a requisite cultural context is missing.

Whether we consider different contexts to be possible worlds or historical episodes of the actual world, it seems that both Danto and Goodman imply that situating the same physical object – or the same syntactic structure – into different environments doesn't change the ontological type of the object in question. What changes from one environment to another are only descriptions of the object. However, if the descriptions change from one environment to another, or one possible world to another, then they are not denoting the given object rigidly. We have seen that both Danto's and Goodman's ontology of a literary work implies that literary descriptions are related to syntactic structures only accidentally. From this it follows that, in the case of literary works we have only texts, or syntactic structures, that function in different ways in different contexts, but literary works themselves are reduced to the status of mere accidental descriptions of these texts. This reductionist conception of the ontology a literary work is unsatisfying in a way similar to the reductionist ontology of the mind.[31]

BIBLIOGRAPHY

Borges, Jorge L. 1962. 'Pierre Menard, Author of the Quixote' in *Ficciones*. New York: Grove Press: 45-56.

Burge, Tyler. 1979. 'Individualism and the Mental' in Peter A. French, T. E. Uehling, Jr. and Howard K. Wettstein (eds.) *Midwest Studies in Philosophy* 4. Minneapolis, University of Minnesota Press: 73-121.

Burge, Tyler. 1986. 'Individualism and Psychology' in *Philosophical Review* 95(1): 3-45.

Burge, Tyler. 1993. 'Mind-Body Supervenience and Explanatory Practice' in John Heil and Alfred Mele (eds.) *Mental Causation*. Oxford: Oxford University Press: 97-120.

Currie, Gregory. 1991. 'Work and Text' in *Mind* 100(3): 325-340.

Danto, Arthur. 1981. *The Transfiguration of the Commonplace*. Cambridge (Mass.): Harvard University Press.

Goodman, Nelson. 1949. 'On Likeness of Meaning' in *Analysis* 10(1): 1-10.

Goodman, Nelson. 1956. 'A World of Individuals' in *The Problem of Individuals: A Symposion*. Notre Dame, Ind.: Notre Dame University Press: 13-21. Page ref. from Paul Benacerraf and Hilary Putnam (eds.) *Philosophy of Mathematics*. New York: Prentice Hall 1964: 197-210.

Goodman, Nelson. 1976. *Languages of Art*. 2nd Ed. Indianapolis: Hackett.

Goodman, Nelson. 1977. 'Worlds of Individuals.' Repr. in Goodman 1984: 48-53.

Goodman, Nelson. 1978a. 'On the Identity of Works of Art.' Repr. in Goodman 1984: 139-142.

Goodman, Nelson. 1978b. *Ways of Worldmaking*. Indianapolis: Hackett.

Goodman, Nelson. 1984. *Of Mind and Other Matters*. Indianapolis: Hackett.

Goodman, Nelson and Catherine Elgin. 1986. 'Interpretation and Identity: Can the Work Survive the World?' in *Reconceptions in Philosophy and Other Arts and Sciences*. Indianapolis: Hackett: 49-65.

Goodman, Nelson and W. V. Quine. 1948. 'Steps Toward a Constructive Nominalism' in *Journal of Symbolic Logic* 12(4): 105-122.

Kripke, Saul. 1980. *Naming and Necessity*. Cambridge (Mass.): Harvard University Press.

Lewis, David. 1983 (orig. 1966). 'An Argument for the Identity Theory' in *Philosophical Papers* 1. Oxford: Oxford University Press 1983: 99-107.

Lewis, David. 1999 (orig. 1972). 'Psychological and Theoretical Identifications' in *Papers on Epistemology and Metaphysics*. Cambridge: Cambridge University Press 1999: 248-261.

Owens, Joseph. 1986. 'The Failure of Lewis's Functionalism' in *Philosophical Quartely* 36(143): 159-173.

Putnam, Hilary. 1975 (orig. 1967). 'The Nature of Mental States' in *Mind, Language, and Reality: Philosophical Papers* 2. Cambridge: Cambridge University Press: 429-440.

Putnam, Hilary. 1981. *Reason, Truth and History*. Cambridge: Cambridge University Press.

Quine, W. V. 1951. 'Two Dogmas of Empiricism' in Quine 1953: 20-46.

Quine, W. V. 1953. *From a Logical Point of View*. 2nd ed. Cambridge (Mass.): Harvard University Press.

Quine, W. V. 1961. *Word and Object*. Cambridge (Mass.): MIT Press.

Tilghman, B. R. 1982. 'Danto and the Ontology of Literature' in *Journal of Aesthetics and Art Criticism* 40(3): 293-299.

Wittgenstein, Ludwig. 1958. *Philosophical Investigations*. 2nd ed. (ed. and tr. G. E. Anscombe). Oxford: Blackwell.

GÖRAN ROSSHOLM

Apropos Menard: A Discussion of the Concept of Literary Work

Abstract: What constitutes an allographic artwork, such as a musical, theatrical or literary work? Several criteria are discussed – such as, same origin, spelling, or expressive qualities – but none is concluded to be sufficient or even necessary. As an alternative, 'hermeneutic equivalence' is considered as a means to define work identity, but again this criterion is found to be inoperative. In conclusion, a difference between literary and other allographic artworks is presented: same- ness of spelling is sufficient for the production of new literary – but only literary – works.
Keywords: ontology of art; works of art; instances of artworks.

Few literary works have caused so much theorizing as Jorge Luis Borges' about ten pages long story 'Pierre Menard, author of Don Quixote', in particular within the fields of aesthetics and the philosophy of language. It provokes ingrained ideas about artistic originality, authentic expression and, most of all, our prejudices about what constitutes a literary work, or more pre- cisely, what constitutes one literary work.

In his classic essay 'The Problem of Meaning in Linguistics' (Quine 1961: 47-64) W.V.O. Quine reduces the number of prob- lems in the philosophy of meaning to two: what is it for a word (or more) to be meaningful, and what is for two words (or more) to mean the same? The two parallel questions about art works run: what constitutes a work of art, and what does it mean to say that two specimens are the same work? The first question concerns how art works are individuated – what counts as one work, what count as two, and what counts as a part of a work. Is the opera Siegfried one work, or is only a third of a work, The Ring? I will take a very liberal stance in this question and not disqualify anything which is usually considered to be only a part of a work from having the status of a complete work. Conse- quently, the text comprising less than three chapters in

Cervantes' Don Quixote, running exactly like Menard's whole Don Quixote, is a work.

The main topic of this article is the second question, about sameness, applied to literature, as it is actualized by Borges' short story. My discussion will also include some statements about literature in comparison to other art forms such as painting, music and theater. I will start by making some – uncontroversial – general claims about singular art works, which will be developed later, and then approach my main topic by discussing some examples from different art forms.

A comma may make a crucial difference between two texts, literary or not. But within certain genres some differences don't count. In a text with the sole purpose to announce a certain truth about identity, the phrase 'x = y' may well be replaced by 'y = x' without any substantial change. However, in a textbook in logic, where we read '(x = y) Û (y = x)', such a substitution would make the stated proposition pointless. But, on the other hand, in this book we may be free to replace a full stop by a comma followed by the copula 'and'. In literature, any change of spelling makes a difference. The same goes for music: any sound emitted from the orchestra makes a difference. In pictorial art we may go still further: any difference may count. Usually only such properties which are detectable by eyesight are considered to be relevant. But that is not always the case. The American artist Jaspers Johns' bronze sculpture Painted Bronze (Ballantine Ale) from 1960 represents two beer cans. It differs from what it represents by possessing some properties not shared by any real beer cans, among them its weight. The fact that the sculpture is much heavier than the represented objects contributes to its meaning. It is heavy, literally and figuratively, and its literal weight is a means to express its metaphorical heaviness.[1]

In short, the reception of aesthetic objects and events seizes upon many more perceptible properties than the reading of watches, maps, thermometers and other non-aesthetic symbols. If we call this characteristic of the arts the Everything Counts Principle we should keep in mind that this slogan-like name exaggerates the range of relevant properties and also that it fails to make any distinction among the various forms of art. I will return to these two questions at the end of this article.

---

[1] By 'expression' and 'expresssivity' I mean, like Goodman (1976: ch. II), metaphorical exemplification.

The question whether two things are instances of one and the same work is meaningful only for what Nelson Goodman (1976: 113-122) calls 'allographic' art, not for 'autographic' art. A painting is autographic, and as no individual object is identical to any other individual object, no painting could possibly be the same as another, however similar. Two same-spelled texts could however both be instances of the same work, for example Don Quixote. (In this article I will sometimes use 'instance of a work' and 'work' synonymously; the context will protect from misunderstanding.) Goodman draws the line between autographic and allographic art by one criterion: the possibility of forgery. A painting, such as Leonardo's Mona Lisa, but not a text, such as Don Quixote, or a piece of music, such as Beethoven's Ninth Symphony, may be forged. If someone paints brush-stroke for brush-stroke like Leonardo, and further claims that the resulting work is Leonardo's Mona Lisa (which we assume has disappeared from the Louvre), then he has produced a forgery. But if a composer or a group of musicians would do the same thing with respect to Beethoven's Ninth, they would simply be telling the truth. And this goes for literature too. If Pierre Menard – contrary to Borges' story – claimed that he had produced an instance of three chapters of Miguel de Cervantes Saavedra's Don Quixote we would accept his words as true. Allographic art may be involved in deceitful transactions of many kinds, it may be intentionally distorted and it may be plagiarized – but it may not be forged.[2]

Distortion or not is a central question when critics and scholars discuss work identity. Robert Wilson inserted a well-known Swedish Christmas- and drinking-song ('Hej tomtegubbar') in his production of August Strindberg's A Dream Play (Stadsteatern, Stockholm, 2000).[3] Does this change make it another play, not belonging to the class of instances of Strindberg's drama? The song itself was sung in accordance to the traditional score, and the lyrics were the right ones. But the tempo and the expression of feeling were far from the traditional way of executing this piece. One line – "we live here for a short time, with much labor

---

[2] Cervantes' manuscript may of course be forged, but a manuscript as manuscript is an autographic work. A 'new' novel of Cervantes may also be produced, but that is another kind of forgery – what Umberto Eco (1990) calls 'Forgery ex nihilo'.

[3] The title of the song is untranslatable; a literal but still unsatisfactory English attempt is 'Hey goblins!'.

and much pain" – received more emphasis, and the song was sung in a very slow tempo, and in a very elegiac mood. Do these changes of emphasis, tempo and expression make it a different song? A third example. Ingmar Bergman himself made the translation of Henrik Ibsen's Ghosts for his production of the play at Dramatiska teatern in Stockholm, 2002. In the original Osvald suffers from syphilis, a fact that is never explicitly stated, but which is still obvious for the audience. Bergman changed this disease into aids; this is never explicitly stated but is all the same clear to everyone in the audience. (Osvald suffers from a 'blood disease' in Bergman's version.)

The three examples differ. Wilson's production was played in Swedish, and almost all the words written by Strindberg were uttered by the actors performing in the roles. Bergman's text is a translation from the Norwegian original, and it departs consequently from the lines Ibsen wrote. However, as the Norwegian and the Swedish languages are closely related, it is possible to produce a Swedish translation, which is, roughly, semantically equivalent to the original. But both Wilson's and Bergman's productions contain words never written by Strindberg and Ibsen respectively. This has semantic consequences. For instance, the actor playing Osvald represents a man who suffers from an illness never heard of by Ibsen. The song – we now disregard its context, the Dream Play – is certainly an instance of 'Hej tomtegubbar' with respect to the words and also with respect to some properties of the score: the pitch and the relative quantity of the tones are preserved from more traditional earlier instances. But the expressive character is totally changed, from the joyful to the melancholic, and the thematic stress is removed from the blessings of drinking to the vanitas theme. Such alterations of mood open to changes of content: the singing men are represented rather as ancient philosophers, Stoics or Epicureans, meditating about the shortness and toughness of life, than as merry Christmas characters.

So far the relation between instances is considered from four different aspects. First, syntactically – the song is spelled, both with regard to words and to tones, in the same way as earlier instances of the same song. Second, semantically, with respect to meaning in a narrow sense, which I suggest we equate to the truth conditions of the representation. The third aspect is what I call expressiveness – what moods, feelings, views etc. are expressed by the work. Finally, we consider origin, the relations

between the performances and what Ibsen and Strindberg actually wrote and intended. In one of the three cases these aspects point in different directions, one away from the work, and one toward it. Osvald's disease is changed by Bergman, who thereby creates a version which deviates syntactically and semantically from what Ibsen wrote. But the change from syphilis to aids may also be said to preserve something expressed by Ibsen's play in the 1880's. Syphilis was a disease associated with sexual excess and unavoidable physical and mental destruction ending in death. Today, when syphilis can be cured, aids has taken over that role. By changing letters and references the new version expresses something more close to what was originally expressed by the work.

There is one more difference between the art works mentioned, between the two plays on the one hand and the song on the other hand. They are all allographic art, that is, instances of one and same work, or distortions of a work, whatever you choose. But though it is possible to catch the song by a notation (or as a matter of fact, two notations), the plays cannot be caught. A notation is – I now follow Nelson Goodman's theory of symbols (Goodman 1976: ch. IV) – a description and a prescription which stands in a one-to-one correlation to segments of a work. Each note in a score matches a tone in the sounding music, and each letter (written letters or phonetic symbols or sounds) in a quotation matches the very same letter (written or uttered) in the work. The symbols in the notation must be discrete: for every mark it must be theoretically possible to determine if it is an instance of that symbol or not. The notation preserves what it refers to: you can go from one instance of a work to a notation to a new instance to a new notation and so on, without any significant change in neither notation nor work.

This only functions partly for drama. The speeches are notational, but the descriptions of scenes or the speech and acting instructions are not. The description 'A brown wardrobe to the left' in A Dream Play may be realized on stage by a brown wardrobe to the left; this object may in the next step be described as 'A dark piece of furniture', which in turn may result in a black sofa to the right on the scene. There is no limit to how far from the original description we can go if we continue along this line.

One way – Goodman's way – to define work identity of notational art forms, such as music, literature and dance, is simply to appeal to the notation. Everything following the score and

quote of 'Hej tomtegubbar' is an instance of that work, and nothing else is. In partly notational art forms such as theatre, conformity to a notation is a necessary but not a sufficient condition. Consequently, Menard's and Cervantes' works (i.e. a fragment of the latter), both entitled Don Quixote, are instances of the same work according to Goodman's theory. In the article 'Interpretation and Identity: Can the Work Survive the World?' Goodman and his co-writer, Catherine Elgin (Elgin – Goodman 1988: 49-65), argue in favor for this conclusion, which is contrary to the position of other philosophical Menard commentators (Lewis 1983; Currie 1990: 42, 77, 78, 178). Works are identified with texts, and texts are individuated with reference to spelling and language:

> Take the word 'cape'. It refers ambiguously – sometimes to an article of clothing, sometimes to a body of land. But it normally does not refer to whatever is either article of clothing or such a body of land. Although ambiguous, 'cape' is a single word. […] In the word 'cape', then, we have a single short text. […]
>   But now consider another equally brief example: 'chat'. This also has two alternative applications: to conversations and to cats. […] But this case is different; for 'chat' is not even the same word in French as in English. […] 'chat' is an English word in 'some chats', but a French word in 'quelques chats'. It is a different word, a different text, in the two languages. (Elgin – Goodman 1988: 58-59)

On the other hand, Goodman's criteria of work identity yield the conclusion that neither Wilson's nor Bergman's productions are performances of A Dream Play or Ghosts. Still harder to accept is the consequence that any deviance, however small, from the score disqualifies the music played from being an instance of the work announced in the program. According to this definition we can't be certain that a work such as Beethoven's Ninth Symphony has ever been performed. If we don't want to accept this conclusion, we must reduce sameness of spelling to a sufficient but not a necessary condition for work identity in notational art forms. However, there may be cases where even this challenges our intuitions about art works. Imagine that two poets, Cain and Abel, independently of each other wrote two same-spelled poems which taken out of context – here assumed to be the two collections of poems – would consist of one ambiguous sentence, which in both books would be disambiguated

by context but carrying widely different senses. According to this weakened form of the syntactical definition they would be instances of one and the same work.

There is one more problem with Goodman and Elgin's position. As seen in the quote above the notion of same-spelling is restricted to texts (or utterances) written (or uttered) in one and the same language; this presupposes that we have to distinguish between natural languages on the one hand and dialects, sociolects and chronolects on the other hand; further, even if it is possible make such a distinction, the restriction to one and the same natural language must be better motivated.

If Goodman's purely syntactical definition of 'same work' is difficult to accept, what about the other three aspects mentioned: meaning in the sense of truth conditions, meaning in the broader sense – what I have called expressivity – and origin?

The narrow meaning criterion has the following consequence: if it is possible to translate Norwegian into Swedish without any distortion of meaning, then the translated text may justly be called an instance of the original work. But we don't have to go deeper into this: it is obvious for several reasons, that identity of meaning in the narrow sense is an inadequate criterion for work identity. First, there are no truth conditions involved in music. And even if we confine ourselves to literature, we confront the same shortcoming: there are literary works without this sort of meaning, nonsense verse, DADA-literature, concrete poetry. Thus: identical truth conditions are not necessary for being instances of one and the same literary work. But neither sameness of meaning in the narrow sense is sufficient for work identity. It is indeed difficult to imagine a set of declarative sentences which would catch the truth conditions of Rimbaud's Le bateau ivre, but if we could, the result would certainly be something else than an instance of this poem. To paraphrase is not to produce a new instance.

The same deficiencies characterize the second category, expressivity. Two different works may both express a feeling of vanity, and several playings of the same music – 'same music' in ordinary parlance – often express different moods.

May we set our hope on the final category? Obviously, having the same origin is not sufficient for two texts to be instances of the same work. We may have different opinions about whether Bergman's and Wilson's productions are instances of A Dream Play and Ghosts respectively, but there are

wild distortions of works, not acceptable as instances of what they distort. Is common origin a necessary condition? Let's take an example. The story of Little Red Riding Hood had existed in many versions before it was codified by the Grimm brothers. Some of these versions were very brute and also obscene, and they ended unhappily.[4] The Grimm brothers are the ones we should praise or blame for the relative oblivion of these earlier versions. But let us imagine another course of events. In a small Bavarian village two storytelling mothers decided independently of each other but influenced by the Zeitgeist to change the shocking ending of Little Red Riding Hood in order to protect their children from fear and evil thoughts. They made up a happy ending including a hunter, who killed the wolf and so on. Independently of each other but by the same reasons they changed the story word by the word in the same way. Their children – who never met – told the story to their children as it had been told to them by their mothers. And so on. Then this modified story reached the four ears of the Grimm brothers. They preferred this version to the more horrid ones also available – and there we are. The question is this: did the creative story-telling women's children and grandchildren listen to instances of one or of two works? Origin says two; common sense isn't so certain. What the Grimm brothers picked up (according to this fantasy) may as well be described as one version, manifested by several instances. Or, to take one step toward Menard: I teach two different courses simultaneously, one about Spanish Renaissance and Baroque Literature and one, on a more advanced level, about Non-Existing Symbolist Writers. I send attachments to my students with literary samples; I scan a page from my own copy of Cervantes' Don Quixote, and use the same in the attachment to the students on the higher level course. Or I send a page from my copy of the Menard edition to the students in the less advanced course, and a page from Cervantes to… Have I deceived the students? Would the dean care? Would anyone? To find one more illustration we may step into Borges' short story. The fictive narrator quotes two lines from Cervantes, which he comments condescendingly. The he quotes the very same lines, but this time 'from' Menard, and he comments them enthusiastically. But what has he done? Ac-

[4] The best known example is found in Charles Perrault's Histoires ou contes du temps passé: avec des moralités (1697).

cording to the narrator, Menard destroyed all his drafts, he burned his notebooks. The narrator never tells us straight out if Menard's final manuscript – consisting of chapter IX, XXVIII and a fragment of chapter XXII, all taken from the first book of Don Quixote – was ever saved. The only glimpses we get from the narrator's reading of Menard show him with an edition of Cervantes, read as written by Menard. We can't rule out the possibility that the narrator has copied the same lines twice from Cervantes' work, and, subsequently, read them differently. In this case it seems as if the source doesn't matter. Use overrides origin.

None of the categories is sufficient, none is necessary. All the same, they are all pertinent to the fictive narrator's reading of Menard, even if they don't play exactly the roles discussed on the previous pages. According to the narrator's interpretation there are differences with respect to truth conditions – Menard's text defines history as the origin of reality, and nothing like that is to be found in Cervantes. There are also differences in style and expressivity – Menard's text is ambiguous and archaic, Cervantes writes an unequivocal everyday contemporary Spanish. But on the whole the distinction between the narrow and the broad semantic aspect is difficult to draw; the differences are better summarized as differences of interpretation. Also the third category, origin, is seen in this perspective in Borges' story. What's important is not the history of the sheets of paper, but how different assumptions about the concrete text's history and prehistory generate and support different interpretations.

This concept of interpretation is very inclusive – traces of influences from Nietzche on Menard and Menard's a bold renewal of the genre of historical novel belong here as well as what happens in the story and what moods are expressed.

Is this somewhat amorfous notion of interpretation what we need to define a literary work? I will come that question after having dealt once more with the category which is the most striking one in Borges' story: sameness of spelling.

Spelling, too, is a matter of interpretation. When I see a certain mark, H, I may interpret it as a letter in the Cyrillic alphabet or as a letter in the Latin alphabet or as no letter at all. If it is framed by a mark, C, to the left, and two marks, AT, to the right I may interpret it as a French or as an English word. This approach – to consider the identification of spelling intergrated

in the interpretation – doesn't lead to Goodman's and Elgin's somewhat arbitrary restriction to natural languages mentioned above. Two texts may be read as same-spelled relative a certain alphabet, and they may also be interpreted as belonging to a certain language.

There were two objections to defining art works by the criterion of spelling: it lets too much in, and it keeps too much out. The reason why it is too wide is that it is a purely syntactical criterion. The reason why it is too narrow is that it doesn't allow for the smallest deviance from the notation, not a half tone too high, not a comma left out. The latter is certainly in conflict with our everyday use of 'literary work'. Different Spanish editions of Don Quixote are certainly not same-spelled, they differ on purpose and by mistake. We could introduce more relaxed claims, not on identity but on similarity. But these claims would all the same have to be built on something like Goodman's idea of notation systems; absolute matching would constitute the extreme on a scale for one condition of work identity. Let us confine ourselves to cases of literary co-instances on which almost everyone would agree, cases which fulfill what we may call the standards of 'strict sameness of instances'. Texts which pass Goodman's syntactical test fulfill one requirement for being strictly the same work, which is of course a subclass of the more vague class of instances of the same work according to ordinary parlance. Instead of trying to define degrees of similarity within the larger class I suggest that like Menard we stick to strict syntactic matching.

The first objection still holds – all we have is a necessary condition for being instances of same work, strictly speaking. The semantics is still missing, so we return to the question about literary interpretation.

At the beginning of this article I claimed that any change of spelling in a work of literature counts. Let me now be more specific.

Not only truth conditions and moods are relevant to the reader of literature. Traces of ideas, allusions to and echoes from other works and traditions, deviation from and adherence to everyday language, to dialects, to foreign languages – all this and much more belong to what the reader could and should recognize in a literary work. This means that every detail in the spelling is relevant. Rhythm, resonance, and variation depends on choices of syntactic construction, uses of homonyms or

near-homonyms, synonyms or near-synonyms, punctuation, number of syllables and much more. (Menard's list of poems, whose effect depends on punctuation marks, bears testimony to this Everything Counts Principle.) And the recognition of any such feature belongs to the interpretation of a literary work.

However, is this really a difference between literature and all other kinds of writing? Einstein wrote '$E = mc^2$', not the equivalent formula '$mc^2 = E$'. Why does it matter? The former but not the latter is the direct and striking answer to the terse question 'What is energy?' But there is no difference between the two formulas with respect to truth or to the rest of the theory of relativity. The difference is more a matter of rhetoric than of physics. Still, it counts.

I suggest that we change the range of application for the literary Everything Counts Principle. It applies when we read texts aesthetically, as art, irrespectively of what we read. Such readings are called for by literature, but not only by literature.

One specification of the literary EC Principle is left: what does not count. When I read Le bateau ivre in a copy printed in Times New Roman, on a page stained by drops of coffee, neither of these circumstances counts. They may influence my reading but I know that they are irrelevant. The EC Principle means that I should attend to certain perceptible features shared by the concrete inscription I hold in my hand and any other instance – in the strict sense – of the same work. This means in turn that my interpretation is an interpretation of every instance of the work. It also means that I only have to seize upon these features if I want to produce a new instance of the work. If we call the properties necessary and sufficient for a proper replication (i.e. a new instance of a certain work of art) the replication qualities of the work, we may say that the replication qualities for literature are the spelling, nothing more. This is not to embrace Goodman's syntactical definition of literary works. Cain's and Abel's poems are not written in order to be instances of the same work, and Menard's Don Quixote is not intended as just another copy of a part of Cervantes' novel. To define strict aesthetic identity we should consider other aspects of interpretation as well.

As said above an interpretation of one instance is an interpretation of every instance of the same work. I suggest that being an instance of one and same literary work in the strict sense could be equated to being mutually indiscernible from a literary point of

view. The instances may be located in different bookcases; they may be printed in different ways and so on, but in every aspect that has to do with the texts as literary texts they are indiscernible. The spelling must be the same and the literary interpretations must be the same. This does not mean that every new interpretation creates a new work. Instead it means this: if the interpretation is valid, it is valid for every instance. If it is awkward, it is so with respect to every instance. If an interpretation is intentional, it is intentional with respect to every instance. My reasons for interpreting the instance as I do is as good (or bad) as they would be for interpreting any other instance of the work in the same way. What kind of interpretation I come up with doesn't matter as long as it is a literary interpretation. A twisted and obscure Lacanian Postcolonial Gender interpretation is twisted and obscure with respect to every instance of the work. It is neither required that different interpretations of instances of one and the same work should be mutually consistent. All that matters is that each interpretation applies to – is about – every instance and that every significant relation between the interpretation and the instance also holds between the interpretation and every other instance of the same work.

Of course, one particular interpretation may comprise more than the instances of one work – I may interpret two poems as sad or as being spelled so-and-so. To narrow the field of application I propose that we incorporate the objective reasons for every interpretation into what I will call 'the principle of hermeneutic equivalence'. I suggest this principle may be spelled out as follows: text A and text B are instances of the same work if and only if there is no objective reason to interpret A differently from B.[5]

What are the consequences of this idea? Cain's and Abel's same-spelled but far from synonymous poems are different works; the interpreter interprets the text in Cain's book differently from the one in Abel's book. This probably accords with

---

[5] There are a few works of literature which cause some trouble to the suggested explication. In his novel Rayuela the writer, Julio Cortázar, tells us that his chapters may be properly read in any order. If a reader starts in the middle, reads to the end and finally reads backwards from the middle to the beginning of the book he may justly claim that his interpretation is valid only for such readings of the book, not for every reading of it. To cover this case we must reformulate the claim on hermeneutic equivalence: text A and text B are instances of the same work if and only if there is no objective reason to interpret A differently from B when read in the same order.

our intuitions. If the same-spelled modified versions of Little Red Riding Hood with two different origins are interpreted identically they are instances of the same work, later to be picked up by the Grimm Brothers. This also seems okay.

What about the two students? If I put a page of a Cervantes edition into the scanner, will the student studying Spanish Renaissance and Baroque Literature receive an instance of a page by Cervantes and the student on the more advanced level a Menard page? Does what technically speaking seems to be an instance Cervantes' Don Quixote change its identity in the attachment sent to the more advanced student? And if the two students are partners, living and studying together in one apartment with one computer, they read the same concrete text. What text? Well, if the advanced student reads first and the other after, maybe we can talk about instances during certain times? From 12 noon to 1 p.m. the text is an instance of a page by Menard, from 1 p.m. to 6:45 p.m. the less advanced student reads an instance of a Cervantes page. Does this mean that we cannot always identify work instances with enduring objects, but sometimes only with temporal parts of objects? Such a chameleon solution has probably little support by our pre-theoretical ideas about work identity. Still worse, if the students read simultaneously, what then are they reading? It can't be a text which is both an instance of Menard page and an instance of a Cervantes page. The idea about hermeneutic equivalence would in that case imply that all Menard instances and Cervantes instances are of one and the same work, which is at the same time contradicted by the differences in interpretation. The relation between two instances of the same work is strong; it is reflexive, symmetrical and transitive. In this case it seems as if we are forced to the conclusion that what appears on the screen is neither an instance of Menard's nor of Cervantes' Don Quixote.

I suggest that the principle of hermeneutic equivalence gives some guidance in these cases. To read the text (copied from an edition of Cervantes' Don Quixote) as written by Menard means to read it as written in an archaic style. The situation – the student's participation in the course Non-Existing Symbolist Writers – motivates this reading. But the student, knowing the text being taken from a Cervantes edition, doesn't recognize the text to be written in an archaic style. If we consider such recognition to belong to the interpretation, the principle of hermeneutic equivalence favors the conclusion that both students read an

instance of one and the same work, i.e. a page of Cervantes' Don Quixote. As already said, what matters in this case is the use of the text. But this only means that sometimes the identity of the work is of no practical importance.

Generally speaking, if literary interpretations are given the form 'The reader reads the texts as …' this allows for three alternative specifications: the interpretation is factual whenever the reader presupposes that the correlated sentence 'The text is …' is true; the interpretation is counterfactual whenever the correlated sentence is false; the interpretation is afactual whenever the reader doesn't presuppose anything about the truth-value of the correlated sentence. I suggest that we consider these presuppositions to belong to the interpretations (cf. Rossholm 2003). Consequently, there are objective reasons to interpret the concrete text descending from Cervantes' head and hand differently from any instance of Menards work. Further, any instance of Cervantes' work may justly be given a Menardian interpretation, but this interpretation will be a counterfactual one.

This is not to equate work to origin. The two story-telling mothers are still producing one work. To say that a reader reads x as invented by Frau A, not Frau B, and that this a factual interpretation, is no literary interpretation if nothing more follows from it.

However, the principle of hermeneutic equivalence has two obvious and serious shortcomings. It might be the case that there is some reason, unknown to us, to interpret the two mother's same-spelled stories differently. There might be some relevant facts in the histories and pre-histories of the instances, which motivate different interpretations, and we can never prove that this is not the case. Further, the principle doesn't help us to determine whether two inscriptions are co-instances of the same work. The idea about hermeneutic equivalence has no operational value for defining a literary work. We cannot ask readers to interpret concrete texts spelled exactly as Cain's and Abel's poems hoping to find out which texts are instances of the one and the other work. Rather, the principle reflects the reader's attitude toward concrete texts which are already recognized as instances of particular works.

Even if neither Goodman's criterion of spelling nor my proposal about hermeneutic equivalence may serve as a satisfactory definition of literary work in the strict sense, they both cast

light upon a difference between literature and other forms of allographic art.

The reader of literature is committed to consider his or her interpretation of a concrete text valid for every other instance (strictly speaking) of the same work. (That is what is left of the principle of hermeneutic equivalence.) And, we may add, he or she may reasonbly assume that there are such co-instances, or at least that such instances may be produced.

What about music, dance and theatre?

Instances of music are produced by playing; instances of dramatic art are produced by performing. But different playings of the same musical work must certainly not be interpreted identically by one and the same interpreter on one occasion. The interpreter might consider a traditional version of 'Hej tomtegubbar' captivating and elevating and Wilson's same-spelled version of the same piece sad and depressing. The same with theatre performances. In these cases the instances are themselves aesthetic interpretations, contrary to the production of books.

This means that we are never committed to consider our interpretation of the performance we watch or listen to valid for any other instance of the same work not because co-instances in the strict sense are held to hermeneutically unequivalent but because we have no reason to believe that there are any such things as strict co-instances in the performance arts. Any two versions of any piece of music will differ significantly. Even if they may both follow the same score, they will still differ in other musically significant respects – in intensity, in modulation, in phrasing, in tempo, that is in respects which are not notational, but which are still aesthetically important. The same is true of dance, and of theatre performances.

This unique position of literature among the fine arts calls for a specification. Two recitals of the Iliad by two rhapsodes may certainly be interpreted differently – understood and evaluated differently. Literature is – in conformity to the etymology of the word – written, by hand or machine.[6] Oral 'literature' is another art form. However, there are art works usually considered to be pieces of written literature, which still don't seem to fit. In several poems by Apollinaire (his 'calligrams') the design and the

---

[6] This fact calls for more details in the example with Little Red Riding Hood: the Bavarian mothers, their children and grandchildren wrote down the stories they invented and listened to.

distribution of the letters bring qualities to the work, which are relevant to the interpretation, but which can't be captured in any notation. However, such works are no counterexamples to the proposed sense of 'literary work' – they are autographic art, without co-instances, like paintings and sculptures. Hermann Hesse's Der Steppenwolf may exemplify another kind of written literature. In most editions the part called 'Traktat vom Steppenwolf' is printed differently from the rest of the text, in smaller letters, or in another typeface. However, these printing properties are certainly notational. The notion of spelling can easily be extended to harbor discrete differences of size, distinct typefaces, italics/plain text and many other typographical categories. The question is what we shall count. Is the choice of relative size of letters constitutive of the work, or not? But whether we answer Yes or No is of no importance for the classification of the work as notational.

This leads to the question what makes literary works (or works read as literature) unique in the art family? It can't be that they are allographic or notational or that they belong to an art. So are music and dance and, partly, drama. Neither that they are verbal – so is oral 'literature'.

I suggest that the difference between literature and the other allographic arts hinges upon from their different replication qualities. As mentioned, I can produce a new instance – in the strict sense – by copying a text letter by letter but I can't create a new instance – in the strict sense – by just following the score or the dance notation or the dramatic script. The replication qualities in literature and only in literature are notational in Goodman's sense.[7]

What about sameness of work in a less strict sense? We may develop a coordinate system with two values: one for deviancies of the notational properties and one for the rest of what we consider might belong to aesthetic interpretation. Strict sameness takes the value zero on both scales, and different objects and events may be located with respect to one another within this

---

[7] The uniqueness of literature cannot be explained by the fact that notation of a work of literature may be conceived to be a quote of the text, i.e. an instance of the work. This is not true of musical score and dance notations: the quote is in itself an instance of the work, but a score is no sounding music and a dance notation moves no limbs. But this difference is superficial. Music may as well be symbolized by played musical quotations, which are instances of the musical work. All we need is the knowledge how to read them.

two-dimensional space. But this prospect is too simple: Bergman's choice of aids instead of syphilis in his production of Ghosts departs from Ibsen's letter, but also from the interpretation that Osvald is represented as suffering from syphilis. On the other hand it is closer to Ibsen's assumed intention that Osvald should suffer from a disease associated with sexual excess and unavoidable physical and mental destruction ending in death. Consequently, it should be dotted at two positions on one of the axes. And since there is no definite number of independent categories relevant for judgements of similarity between instances (i.e. relevant for aesthetic interpretation) and no obvious way to balance conflicting values every attempt to define an overall scale of degrees of sameness is doomed to fail. Like 'strict sameness of instances' the concept of work identity in general is vague. But while the vagueness of the former is restricted to rather marginal cases (like the example with two students), it is characteristic of the whole field of application of the latter.

It does not follow from this that there are two works entitled Don Quixote, one by Cervantes and the other by Menard, not even in Borges' short story. In the discussion above about the two students the crucial question was whether a certain concrete text descended from Cervantes' original or from Menard's. The concept of literary descent was taken for granted. However, Borges story challenges this concept as well as the idea about same-spelling. Gregory Currie (1990: 178, 42) talks about Menard's "(relative) independence from the original text", and states that "if Menard had simply copied out Cervantes's text, he would not have produced any work". But what does relative (within parenthesis) independence and simply copied mean? Menard obviously intended to write a text letter by letter spelled like Cervantes' Don Quixote. Isn't that an intention to copy? It is true that he – according to the narrator – didn't copy Cervantes' novel mechanically, or in any ordinary way. But why not just say that he copied it in an extraordinary way? It is obvious that the act of creation (or copying) was extraordinary. The narrator describes Menard's project as follows:

> Pierre Menard did not want to compose another Quixote, which surely is easy enough – he wanted to compose the Quixote. Nor, surely, need one be obliged to note that his goal was never a mechanical transcription of the original; he had no intention of copying it. His admirable ambition was to ptoduce a number of

pages which coincided – word for word and line for line – with those of Miguel de Cervantes.

The act of creation – or copying – becomes still more obscure when we read Menard's own words about his reading of Don Quixote:

> At the age of twelve or thirteen I read it – perhaps read it cover to cover, I cannot recall. Since then, I have carefully reread certain chapters, those which, at least for the moment, I shall not attempt.

And, as already said, it is far from obvious that anything of Menard's quixotic remains was ever left. When the narrator reads a chapter which Menard never grappled with, he recognizes 'the style of our friend, and something of his voice'. This doesn't create any now works, just new interpretations. As long as we don't have any concrete instances we don't have a work. And as Borges' story goes along, the emphasis is shifted from writing to reading. In the final lines the narrator finds literature enriched by the possibility to read Thomas à Kempis' De Imitatione Christi as written by James Joyce or by Louis-Ferdinand Céline. These recommendations don't create new works, just new interpretations that together with the old ones apply to the same old instances of the same old work.

BIBLIOGRAPHY

Currie, Gregory. 1990. *The Nature of Fiction*. Cambridge: Cambridge University Press.

Eco, Umberto. 1990. *The Limits of Interpretation*. Bloomington: Indiana University Press.

Elgin, Catherine and Nelson Goodman. 1988. *Reconceptions in Philosophy and Other Arts and Sciences*. London: Routledge.

Goodman, Nelson. 1976. *Languages of Art: An Approach to a Theory of Symbols*. Indianapolis: Hackett.

Lewis, David. 1978. 'Truth in Fiction' in *American Philosophical Quarterly* 15: 37-46.

Rossholm, Göran. 'The Tree of Interpretation' in Carlshamre, Staffan and Anders Pettersson (eds.) *Types of Interpretation in the Aesthetic Disciplines*. Georgetown: McGill-Queen's University Press, 2003.

Quine, W.V.O. 1961. *From a Logical Point of View*. 2nd ed. Cambridge (Mass.): Harvard University Press, 1953.

PETR KOŤÁTKO

Text, Work, Interpretation: Some Implications of the Menard Case

Abstract: The paper distinguishes three levels of the text, specifies relations between their identity conditions, and argues that literary interpretation includes two mutually dependent moves: identification of the highest (most complex) level of the text and identification of the literary work. Both moves essentially depend on extratextual sources (including references to the author's intentions), as the Menard case convincingly shows. The extratextual basis of interpretation can be approached either as constituted by interpreter-independent facts of various kinds (the relative weight of which is a matter of continuous discussion), or as a field for the interpreter's constructions, limited only by a specific (literary) version of the principle of charity. The latter approach is exemplified by the "new art of reading" advertised at the end of Borges' story.
Keywords: text-work distinction; context dependence; authorial intentions.

By the Menard case I mean a situation in which we are faced with two type-identical texts, created in radically different circumstances, and the question arises whether they represent the same literary work. Analogies in music and fine art are easy to see: in the former case we consider two type-identical scores, in the latter two visually indistinguishable physical objects. While some authors inspired by Borges have construed their own examples of this kind,[1] I will remain true to Menard, due to my poor imagination, but will not feel obliged to preserve all the details of the Borges version: in particular, I will suppose that, unlike in the story, Menard has succeeded in completing his *Quixote*. The question is how to apply, in the description of this situation, notions like text, work, content, context and intention.

---

[1] For example, Gregory Currie (1991: 328-330) discusses a fictional case of two different literary works with the same text (unlike in Borges' version, the text-identity is here a matter of accident); Jerrold Levinson (1990: 70-72) construes several examples of clearly different compositions with the same score; Arthur C. Danto (1981: 1-2, 39) describes a series of visually indistinguishable paintings representing different works of art.

As far as the first notion is concerned, I am afraid that further specifications will be unavoidable. For our purposes we ought to distinguish (at least) the following:

$Text_1$: a syntactically identified series of symbols (types);[2]

$Text_2$: a syntactically identified series of symbols (types) interpreted according to the semantic rules fixed by the conventions of some language;[3]

$Text_3$: a syntactically identified series of symbols (types) interpreted as used to communicate certain propositional contents with certain illocutionary forces.[4]

Sometimes it will be useful to speak simply about a text, taking $text_1$, $text_2$ and $text_3$ as its hierarchically ordered levels.

Moreover, the text (in any of the senses just mentioned) ought to be distinguished from the Literary work: a structured complex of literary (poetic, prosaic, dramatic) aspirations of the text and the literary qualities fulfilling (or satisfying) them – if they are fulfilled.

---

[2] By Cervantes' $text_1$ of Don Quixote I mean a set of $text_1$ instances, one of them being Cervantes' original manuscript. This instance is privileged in the sense that anything else belongs to the set only if it is linked with Cervantes' manuscript by the chain of copying or reproducing. This is a direct analogy of Kripkean causal chains linking utterances of proper names with the original acts of baptism (cf. Kripke 1972). Just as two phonologically (or typographically) type-identical utterances of a name can belong to two different Kripkean chains, and hence differ in reference, two type identical sequences of sentences can belong to two different reproduction chains, one of which begins e.g. with Cervantes' manuscript and the other e.g. with Menard's manuscript.

[3] As to the relation between $text_1$ and $text_2$, we should keep in mind that the syntactic identification of expressions cannot be isolated from their semantic interpretation: those phonetic or typographic distinctions which serve (in given language) to indicate semantic distinctions are exactly those that are syntactically relevant.

[4] In David Kaplan's terms (Kaplan 1989), the step from $text_2$ to $text_3$ is a transition from 'characters' to 'contents'. Needless to stress, conventional meanings of natural languages' expressions do not precede their use in communicating propositional contents with illocutionary forces: on the contrary, the conventional meanings are established and fixed precisely in this sphere. This note and the preceding one are intended to stress that the series $text_1$ – $text_2$ – $text_3$ is not to be interpreted as a straightforward development from more elementary to more complex phenomena .

Although I cannot provide any non-circular definition of literary aspirations, here are at least some examples: the aspiration to conform to conventions of certain genre; to preserve a certain rhythmical pattern; to tell a thrilling story; to evoke the atmosphere of a late afternoon on a deserted country square; to let the character *X* put keys on the table during his conversation with the character *Y*; to refer implicitly to some passage of the New Testament. One literary aspiration can imply another, or represent a specification of another; an aspiration may be indicated without being fulfilled and this discrepancy can play an important role in the fulfillment of another aspiration and hence in the literary construction of the work (cf. the case of parody, creation and frustration of the genre-expectations and similar kinds of plays with the reader); the same is true for conflicts or tensions between incompatible aspirations indicated in the work, etc. One and the same feature may be ascribed to the text in two different ways, depending on our identification of its literary aspirations: as their fulfillment or as a failure. The term 'aspiration', as used here, is to be understood in an equally impersonal way as the term 'intentio' in 'intentio operis', i.e. as a parameter of the work. The relation between aspirations of a work and the intentions of its empirical author is to be left open at this level. (For the term 'literary aspirations' as a correlate of the semantical term 'truth conditions' or more generally of 'satisfaction conditions', cf. Koťátko 2004.)

1. TEXT₁ AND TEXT₂

Let's now consider a possible world in which Cervantes has written a text$_1$ entitled *El ingenioso hidalgo Don Quijote de la Mancha* and about three hundred years later a man called Pierre Menard, in circumstances described in the Borges story, has written a text$_1$, which is sentence by sentence type-identical with the former.[5] The first question is whether they represent the same text$_2$. Obviously, this will be the case if we apply the same semantic rules to both texts$_1$. In Cervantes' case these rules must be conventions of Spanish of his time (and the only

---

[5] This way of introducing the problem does not leave any space for the objection that we fall into confusions by comparing a real work of a real author with a fictitious work of a fictitious author; cf. Wreen 1990. The presumption that two texts$_1$ are type-identical evokes the problem of the criteria of syntactic identity, which while far from trivial (cf. Russell 1961: ch. 2), does not belong to our present context.

problem concerns the completeness and reliability of our knowledge of them). But in Menard's case we face a dilemma: should we assign to the expressions contained in his text₁ [a] the meanings which (as far as we know) they had in Spanish (Castilian) of the beginning of the 17th century; or rather [b] the meanings which (as far as we know) Menard or the experts of his time thought to satisfy [a]?

Suppose we find out that [a] and [b] in some particular cases differ and that these differences are relevant for literary interpretation. Nevertheless, we may feel inclined to ignore this discrepancy and opt for [a]. As we are given to understand from Borges, Menard intended his text₁ to be read as written in the Spanish of Cervantes' time (and hence to be interpreted according to its rules). And Menard's text₁ indicates this intention quite strikingly by its type identity with Cervantes' text₁. But there are also reasons speaking clearly for [b]. First, we know that Menard's *specific* semantic intentions (concerning the interpretation of particular expressions) in some cases deviate from his *general* semantic intention (to write in the Spanish of Cervantes' time). To interpret his text₁ according to [a] then means to assign deliberately to some expressions meanings which differ from the meanings assigned to them by Menard (as well as by the experts and competent readers among his contemporaries). To abstract from this discrepancy is to approach Menard's general semantic intention as self-satisfying, in the sense of being fulfilled by being recognized. But this creates an undesirable precedent: in the same way we should then approach Menard's general literary aspiration, i.e. his intention to create a literary work whose text will be identical (on the level of text₁ and text₂) with that of Cervantes, but which will nevertheless represent an accomplishment of an original literary project, anchored in (and interpretable from) the context of the beginning of 20th century. Menard's achievement, i.e. his working out all the details of that project and its realization in the textual form, would then be redundant.

I do not want to suggest that this is a particularly interesting dilemma from the theoretical point of view. Most of the following discussion will focus on the relation between text₂, text₃ and the work: the identity of Cervantes' and Menard's texts₂ will be taken as granted. My pedantic remarks were just meant to show that this identity is not guaranteed by the description of

Menard's case in Borges' story. However fixing it by fiat is certainly a legitimate move within our thought experiment.

2. TEXT₂, TEXT3, WORK, CONTEXT
2.1. LITERAL MEANING

The identity on the level of texts₂ does not guarantee the identity on the level of texts3: this should be clear from an analogy with ordinary communication. Grasping the proposition expressed in a particular utterance of the sentence 'This is John's bank' requires more than knowledge of the conventional meaning of the sentence (its 'character', to use David Kaplan's illuminating term): contextual knowledge is needed for an identification of the referent of the demonstrative, for the choice of the relevant meaning of the homonym 'bank', and of the relevant reading of the apostrophe, as well as for picking out the right man among thousands of bearers of the name 'John'. In general, we have to fix those parameters of the propositional content (and similarly of the illocutionary force) of the utterance which are not determined by the application of linguistic conventions on the sentence uttered.[6] If we face a similar task in reading a novel, the source from which we get material for filling in these gaps is the nearest or broader fictional context of the utterance: everything which can contribute to the identification of the speaker's (here: narrator's or character's) communicative intentions. We have to exploit relevant parts of our knowledge concerning the speaker's non-communicative intentions, beliefs, preferences, education, idiolect, conversational skills, specific sense for continuity etc. In other words, even if we remain in the sphere of literal meanings, and if we read the words uttered in conventional way, the transition from the text₂ to the text3 is impossible without identification of certain elements of the literary work.

Moreover, these elements need not be determined on the level of the text₂ and those parts of the text3 which have been already fixed. Let us call a text which is fully determined on the text3 level by what is provided on the text₂ level 'closed' or 'saturated'. I do not intend to claim that the notion of a closed

---

[6] As to the conventionally fixed meaning of the words uttered, let us presuppose (for the sake of simplicity) that we have good reasons to transfer them directly into the meaning of the utterance. Obviously, there are countless situations in which this would lead to misunderstanding, cf. the discussion of the phenomenon of malapropisms in Davidson 1986.

text is incoherent or cannot have any interesting literary exemplifications: I just want to point out that there is no reason to approach *a priori* any literary text as closed. The closeness would be guaranteed only if the $text_2$ satisfied quite strong constraints, like: it does not include indexicals, unless their reference is additionally fixed in a context-independent way (which makes the indexical reference redundant); homonyms may not appear without supplements that explicitly identify the relevant meanings; each sentence includes an explicit performative formula, determining unambiguously the illocutionary force of its utterance.[7] In other words, it would strongly differ from what we know from everyday communication as well as from literary narration.

In case that the $text_2$ underdetermines the $text_3$, so that the transition from the former to the latter requires a broader basis provided by the literary work, the question arises as to the source of those parameters of the literary work that are not textually fixed. Obviously, they must originate in some complex of assumptions which represent a legitimate source of literary interpretation without being justified by reference to the text. A candidate that naturally suggests itself is the complex of beliefs shared by the author and the competent readers of her time. In order to specify their role in the identification of the literary work, one can apply the well-known notion of *truth in fiction*, as introduced by David Lewis. Leaving details of Lewis' definition aside,[8] the application of his notion for our purposes can be described as follows: Let us have some proposition *p* such that both its truthfulness and its falsity is compatible with what is explicitly said in Cervantes' *Don Quixote* (which means: with those elements of Cervantes' $text_3$ that are fully determined by his $text_2$). The question is whether *p* is true or false in Cervantes' novel (or, if you wish: in the world of his *Don Quixote*). In order to see this, let us consider the set of possible worlds in which the story of the novel is told not as a fiction but rather as a fact known to the speaker. Among those worlds, let us pick out the one which is closest to the belief world of the community of the novel's origin – the world *w* which most closely corresponds to the beliefs shared in the community in

---

[7] The last condition is unrealistic: as Davidson (1984) has argued, no (linguistic or non-linguistic) device can in fact have such a power.

[8] In fact, I exploit here one of the two versions offered by D. Lewis (1983: 273).

which Cervantes' *Quixote* was (actually) written. Then if *p* is true in the world *w*, it is true in Cervantes' novel; if it is false in *w*, it is false in the novel.[9]

One can expect that the shared beliefs intervening in this way into Cervantes' novel and its text3 and the beliefs that do this job for Menard differ in many respects. Although this makes room for a variety of differences between both literary works as well as both texts3, I will confine myself to one special case: differences (on the level of texts3) in the function of the first person pronouns, as used to refer to the narrator.

The narrator is a literary construct, a fictional person characterized by his or her vocabulary (and its implications concerning his or her social position, education, temperament etc.), the repertoire of narrative figures, specific ways of addressing the reader, sympathies and antipathies, ideological inclinations, presumptions concerning what has to be said explicitly and what can be presupposed as a shared basis of communication, etc. If we take it that Cervantes' narrator is construed as Cervantes' contemporary, while Menard's narrator is construed as Menard's contemporary, presenting himself as a person living at the beginning of 17th century, the reference of the first person pronouns is obviously different. However, let us make the more natural assumption that both Cervantes' and Menard's narrator is construed as Cervantes' contemporary. In Cervantes' case, this construction is based on his acquaintance with his sociocultural environment, while Menard builds on hypotheses, based on the picture of Spain of the beginning of the 17th century available in France at the beginning of the 20th century. It is then natural to expect that the construction of the narrator will differ in respects relevant for literary assesment, in particular in the (above mentioned) implications of his vocabulary and in the beliefs, preferences and other attitudes which he is supposed to share (and to believe to share) with his contemporaries.

The utterances of the first person pronoun in the *Preface* represent a special case, provided that we read it as not fully inte-

---

[9] It should be stressed that the shared beliefs are not the only elements of the social environment which intervene into the content of fiction: an equally important role play shared dispositions to regard certain behaviour as normal or reasonable and certain ways of speaking as normal or neutral. And deviations from these standards in the behaviour of some characters or in the language of the novel typically create tensions or produce other effects which may belong to the literary construction of the work.

grated into the novel. In case of Cervantes' *Preface* this means that the speaker (and hence the referent of the first person pronouns) is not the narrator but rather Cervantes himself. In case of Menard's *Preface* we have two options:

[a] The speaker is Cervantes as a literary character created by Menard (this is the version held by Borges' narrator) and the novel following the *Preface* is presented as a work of that character (as in those Borges stories in which the narrator presents the narration as a more or less modified transcription of somebody's diary). The first person pronouns in the *Preface* then refer to a literary character, rather than to an empirical person, in contrast to their function in Cervantes' *Preface*.

[b] The speaker is Menard as an empirical person, who (within the *Preface*) presents himself as a writer from the beginning of the 17th century. Then the *Preface* has the same status as in Cervantes' case and consequently the first person pronouns refer in both cases to two different persons (to Cervantes and Menard, respectively).

If, for some reason, we reject the most natural reading and take the *Preface* as an integral part of the novel, its speaker is in both cases identical with the narrator of the novel and Menard's version differs from Cervantes' just in the construction of that narrator (cf. above).

2.2. INDIRECT MEANING, LITERARY FIGURES, INTERTEXTUAL REFERENCES

Until now our focus has been on literal meanings: the question was what is involved in the determination of propositional contents ascribed in literal reading to sentences on the level of the $text_3$. Before we proceed to indirect meanings, we should decide whether we are to include them in the $text_3$ or not. I take this as a terminological matter and suggest to opt for a rather rigid notion of the $text_3$ as including only literal meanings. Consequently, the transition from the literal reading of an occurrence of a sentence in the $text_3$ to an identification of its (possible) indirect meaning will be automatically understood as a shift from identifying the $text_3$ to identifying the work.

In the preceding section we have discussed the importance of extratextual factors for the identification of propositional contents literally expressed in the $text_3$. It is quite evident that once we abandon the restriction to the sphere of literal meanings, the potential role of factors such as implicitly shared beliefs, norms

and dispositions radically increases. In particular, they are needed for filling in narrative gaps, for providing descriptions used in the text with characterizing power,[10] for identifying coherent patterns in the characters' actions and for understanding their motivation, as well as for working out Gricean implicatures (without which we may fail to see the continuity in the characters' conversation and the links between their utterances and their non-linguistic behaviour).

To take an example from the Borges story, let us suppose that we have to choose between a 'serious' or an ironical reading of the narrator's or character's utterance, and that no part of the $text_3$ explicitly specifies their general attitude to the topic of discourse. We can get an idea of which reading to choose from our knowledge (based on the $text_3$) about the narrator's or character's origin, biography, present social position and so on, but only on the background of some presumptions needed for making relevant inferences from such data. Or else we can employ our beliefs about the author's attitude to the matter in question, provided that we have a reason to take the narrator or character as voicing the author's views. This is the line taken in Borges' story: Borges' narrator claims that the adoration of military virtues in the famous monologue of Don Quixote (in chapter 38) should be read as serious in the case of Cervantes and as ironical in the case of Menard ('the contemporary of *La trahison des clercs* and Bertrand Russell'!).

It seems that whatever kind of literary figure we choose, there is a space for potential differences between Cervantes' and Menard's novels: what is stylistically neutral in Cervantes' case, can become an archaism in Menard's case; what works in one case as a powerful metaphor, can in the other be a conventional idiom; what is in the one novel an *epitheton constans*, can in the other be an inventive attribute, etc. The last two modifications can clearly proceed in both directions: hence the Borge-

---

[10] For instance, on the basis of certain presumptions commonly shared in Cervantes' socio-cultural environment things like details of the character's dress or vocabulary can indicate his/her origin or social position and this may generate certain expectations concerning preferences, schemes of behaviour, temperament etc. Against this background, some explicitly described actions may be regarded as natural or surprising and this may be important for understanding their role in the development of the story, for making sense of other characters' reactions or of the narrator's comments.

sian narrator's claim that Menard's *Quixote* is much richer than that of Cervantes (cf. Danto 1981: 35) seems to be one-sided.

Finally, both novels differ radically (though less strikingly) in implicit intertextual references, like quotations (without explicitly specified sources), connotations based on names, motives or typical phrases borrowed from other texts, parodical paraphrases, implicit polemics etc. Some of these intertextual relations, which are transparent and efficient in Cervantes' time, are dead in Menard's time (the relevant texts are simply unknown to Menard and his potential readers, including experts): hence they cannot be a part of Menard's work. And vice versa: Borges' narrator claims to find in Menard's *Quixote* several implicit references to the texts written long after Cervantes' death (cf. above). Hence also in this respect Menard's *Quixote* is both richer and poorer than that of Cervantes.

Everybody will probably agree that identification of the propositional contents of the characters' utterances, identification of indirect meanings (implicatures in Gricean sense), differentiation between neutral and archaic style, detection of the implicit intertextual references etc. are fundamental interpretative achievements: they concern substantial elements of the construction of a literary work. We have seen that these elements are underdetermined by purely textual sources: Menard's case famously points to parameters of the content of literary work which are not internal in the sense of being fully determined by the text (on any of the three levels we have distinguished). We may insist that the literary function of particular elements of the work is given by their place within the structure of the whole. But this whole cannot be limited by the boundaries of the text, just as the whole within which linguistic utterances acquire their meaning cannot be limited by the boundaries of language.[11]

3. TEXTUALISM, ANTI-INTENTIONALISM, CONSTRUCTIVISM
We are now in a position to suggest a reply to the crucial question: what is the proper object of literary interpretation? Within our hierarchy of text levels we should choose the one which is richest in content without depending on the identification of the literary work. This is clearly the text$_2$. Then literary interpretation,

---

[11] Cf. Wittgenstein's claim that 'our talk gets its meaning from the rest of our proceedings' (Wittgenstein 1969: § 229) or the lesson we can take from Davidson's (and his followers') discussion about radical interpretation.

directed to the text$_2$ as its object, consists in identification of the text$_3$ and of the literary work. These two moves, as we have seen in section 2.1, are mutually dependent and underdetermined by the text$_2$.

In general, Menard's case seems to have shown quite clearly that the literary parameters of a work are not essentially internal in sense of being fully determined by the literary text. That means that literary interpretation must have a broader basis than the text in any of the senses distinguished above – and hence that any doctrine identifying text and work (any version of 'textualism', as it is often called) must be wrong.

To be sure, textualists would not accept the line taken in the preceding section. According to them, what we have presented as differences between Cervantes' and Menard's works should be taken as differences between two interpretations of the same work.[12] The problem is that the differences in question concern basic elements of the literary construction of the work, literary qualities *par excellence*: if we abstract from qualities of this kind, we are not left with anything which could be regarded as a literary work to be interpreted.[13] An interpretation assigning to a text archaic style, implicit references to other texts, socioculturally based presumptions which are not explicitly mentioned in the text but are involved in the constitution of the fictional world of the novel and are relevant for understanding motivation and continuity of the characters' actions, for working out implicatures etc., assigns to the text other constitutive literary qualities than an interpretation which differs from the former in any of these respects. One need not subscribe to our account of literary work as a structured complex of literary aspirations and qualities (cf. section I) in order to conclude that the two interpretations in question assign two different works to the same text.

There is yet another argument against textualists, one that appeals to the notion of literary work as a result of an original

---

[12] Cf. Elgin – Goodman 1988: 63: "Menard may in some way have proposed or inspired a new interpretation of the text. But no more than any other admissible interpretation offered before or since or by others does the Menard reading count as the work Don Q., or even a work Don Q. All are merely interpretations of the work."

[13] This contrasts with Elgin and Goodman's requirement (Elgin – Goodman 1988: 56) that the notion of literary work admits the possibility of different (literary) interpretations of the same work.

creative act. If we identify work with text taken as an abstract entity and if we grant timeless existence to abstract entities, we give up the possibility of approaching the empirical author as a creator of the work. The author's achievement then consists merely in producing a publicly accessible instance of the work, and thereby attracting our attention to the work. Jerrold Levinson (1990: 66–70) takes this to be a decisive argument against any account of a work of art as an abstract entity, insofar as one of the crucial premises of our way of thinking about art is that it is a sphere of creative activity producing something which otherwise would not come to being.

Instead of discussing this argument in detail, let us note that both accounts of the author's achievement are compatible with the results of our preceding discussion. We only have to insist that the work, no matter whether created or 'revealed' by the author, is not fully fixed by the text. It is only the textualist account of the work as an entity *fully identifiable by the text* which implies that Menard – independently of his intentions and of the socio-cultural context – produced an instance of the same novel as Cervantes did three hundred years before him. One can be the most radical Platonist and still insist that an identification of the novel that was 'revealed' in Menards' achievement (i.e. the selection of the right complex of literary qualities among all those compatible with Menard's text$_2$) is impossible without appeal to facts concerning the empirical author and his socio-cultural environment.

It is a matter of debate between intentionalists and conventionalists what the broad (extratextual) basis of interpretation should include. In the preceding paragraph we have had several opportunities to appreciate the relevance of the socio-cultural context of the text's origin for the transition from the level of text$_2$ to the level of text$_3$ and to the literary work. However, this cannot be the whole story, as Menard's case also shows. Nothing in the socio-cultural situation of the beginning of the 20th century prevents us from approaching Menard's creative achievement as a mere repetition of Cervantes' achievement (based on empathy), and hence from taking Menard's novel as a duplicate of Cervantes' novel – not to mention the possibility of taking it as mere plagiarism. All the differences between Cervantes' and Menard's works mentioned above make sense only if we presuppose that Menard's aspiration and his method was not like that: namely that he intended to produce a novel

anchored in, and interpretable from, the context of the beginning of the 20th century and that his text2 was a result of his working out this project. Hence we have here a case where an appeal to the empirical author's intentions proves to be of crucial interpretative importance. If this is right, then radical anti-intentionalism, rejecting any reference to the author's intentions as 'intentional fallacy', is untenable as a general strategy of literary interpretation (cf. Danto 1981: 35-36).

As compensation, Borges' narrator mentions another possibility open to interpreters, which we have not yet taken into account: although it may seem rather bizarre, it certainly deserves to be mentioned. To say that literary interpretation must have a broader basis than the text itself does not amount to saying that interpretation requires *knowledge* of certain facts concerning the socio-cultural context of the work's origin and the empirical author's intentions. In other words, the postulated broad basis of interpretation need not be cognitive: it can also be constructed by an interpreter who feels primary commitment to other than cognitive constraints. Most naturally, she may follow a radical version of the principle of charity, saying that she should look for that interpretation which maximizes the text's literary value and that whatever serves to this purpose is good (provided that the principle of centrality of the text is preserved).[14] Or she may, following Richard Rorty (1992), look for a reading which will help her to re-examine and change her personal hierarchy of values. Whatever the ultimate end, this 'new art of reading', based on selecting or construing the context from which the literary text is to be interpreted, was invented by Pierre Menard, as Borges' narrator reminds us in the end of the story.

BIBLIOGRAPHY
Currie, Gregory. 1991. 'Work and Text' in *Mind* 100(399): 326-340.
Currie, Gregory. 1993. 'Interpretation and Objectivity' in *Mind* 102(407): 413-428.
Danto, Arthur C. 1981. *The Transfiguration of the Commonplace*. Cambridge (Mass.): Harvard University Press.

---

[14] Cf. R. Dworkin's formulation of what he calls 'aesthetic hypothesis' (Dworkin 1985: 149): "An interpretation of a piece of literature attempts to show which way of reading (or directing or acting) the text reveals it as the best work of art". Within G. Currie's instrumentalist model of interpreting narration (Currie 1993), we are supposed to construe a hypothetical author as a bearer of such narrative intentions that the interpreted text comes out as their optimal fulfillment.

Davidson, Donald. 1984. 'Communication and Convention' in Davidson, Donald *Inquiries into Truth and Interpretation*. Oxford: Oxford University Press: 265-280.

Davidson, Donald. 1986. 'A Nice Derangement of Epitaphs' in Lepore, Ernest (ed.) *Truth and Interpretation. Perspectives on the Philosophy of Donald Davidson*. Oxford. Blackwell: 433-446.

Dworkin, Richard. 1985. *A Matter of Principle*. Cambridge: Cambridge University Press.

Elgin, Catherine and Nelson Goodman. 1988. *Reconceptions in Philosophy*. London: Routledge.

Kaplan, David. 1989. 'Demonstratives' in Almog, Joseph, Perry, John and Wettstein, Howard (eds.) *Themes from Kaplan*. New York and Oxford: Oxford University Press: 481-563.

Koťátko, Petr. 2004. 'Truth Conditions and Literary Aspirations: Basic Correlations' in Rossholm, Göran (ed.) *Essays on Fiction and Perspective*. Frankfurt am Main: Peter Lang Verlag: 155-170.

Kripke, Saul. 1972. *Naming and Necessity*. Oxford: Blackwell.

Levinson, Jerrold. 1990. *Music, Art and Metaphysics. Essays in Philosophical Aesthetics*. Ithaca and London: Cornell University Press..

Lewis, David. 1983. 'Truth in Fiction' in Lewis, David *Philosophical Papers*, Vol.I. Oxford: Oxford University Press.

Rorty, Richard. 1992. 'The Pragmatist's Progress' in Eco, Umberto *Interpretation and Overinterpretation* (ed. Stefan Collini). Cambridge: Cambridge University Press: 89-108.

Russell, Bertrand. 1961. *Inquiries into Meaning and Truth*. London: George Allen & Unwin Ltd.

Wittgenstein, Ludwig. 1969. *On Certainty* (ed. G.E.M. Anscombe and G.H. von Wright). Oxford: Blackwell.

Wreen, Michael. 1990. 'Once Is Not Enough?' in *British Journal of Aesthetics* 30(2): 149-158.

JAMES HAMILTON
Borges' Menard:
His Work & His Intentions

Abstract: Authors make assumptions about the fitness and adaptability of their
"materials" – such things as multiplicity or simplicity of sentence structures,
techniques for making action appear simple or complex, techniques for achieving
depth or shallowness of character, detailed or spare descriptive language, and so
on – to their aims. But, in this paper I show that this reasonable notion of what
is involved in authorial intentions is not only at odds with, but is actually
obscured by, conceptions of authorial achievement that must be presumed in the
standard discussions of the ontology of art that have employed Borges's story.
Keywords: ontology of art; works of art; adaptability of fit.

In her 1988 essay, 'Losing Your Concepts', Cora Diamond
describes a way in which the adoption of a particular
philosophical theory can render concepts that have been
necessary unavailable for the adequate description and analysis
of important human concerns. Diamond focuses attention on the
way deflationary theories of truth make it impossible to
articulate any sense in which truth itself may be said to be
valuable. And she assembles powerful examples of people who
have claimed that truth is valuable in order to show the
importance of such claims and the impotence of deflationary
theories of truth when faced with the challenge of making sense
of them.

In this essay I maintain that something like this kind of loss
has occurred with respect to the use made of Jorge Luis Borges'
story, 'Pierre Menard, Author of the Quixote', in philosophy of
art. Here, however, it is not a particular theory that renders
important concepts unavailable to us. Instead, I maintain, the
concept of authorial intention – so prominent in Borges' story –
disappears as a robust concept of intention when the story is
taken to raise questions in the ontology of art. It is the focus on
those questions that renders a robust concept of authorial
intention unavailable to us.

To make good on this position, I undertake several tasks. First I sketch the alternative positions on the ontology of artworks that have been thought to be supported or challenged by the several chapters and part of another of Don Quixote that Borges' narrator says Menard produced. Second, I extract from the ontology debate a position on Menard's intentions that is consistent with the positions taken in the debate. Third, in contrast, I sketch an account of authorial intentions derived from attending to details in Borges' story. I also provide grounds for thinking that Borges' implied account of authorial intentions is largely correct. And I show how we should use this account to explain what makes Menard's authorial intentions quixotic. Fourth, I develop an objection to the central claim of this paper, an objection that asserts that we must settle the ontological questions even if we adopt the conception of authorial intentions I have extracted from Borges' story. Finally, I show that any way we have of stating that objection clearly, it falls out that the conception of authorial intentions extracted from Borges' story loses its robustness and gets distorted in such a way that it no longer describes anything like what authors actually do.

1. THE ONTOLOGY DEBATE, AND VIEWS OF MENARD'S ACHIEVEMENT

The stage is set for the use of Borges' story in discussions of the ontology of works of art by Nelson Goodman's claim that

> A literary work…is not the compliance class of a text but the text or script itself. All and only inscriptions and utterances of the text are instances of the work; and identification of the work from instance to instance is ensured by the fact that the text is a character in a notational scheme – in a vocabulary of syntactically disjoint and differentiated symbols. Even replacement of a character in a text by another synonymous character […] yields a different work. (Goodman 1968: 209)

This claim, that work identity is text identity, is what Borges' story has sometimes been taken to challenge.

The challenge has been made in at least three, equally unsuccessful, ways. The first challenge, at least historically, is found in an essay by Anthony Savile (Savile 1971) wherein Savile uses an account of the story to contest Nelson Goodman's view that notational or syntactical identity of a text

alone is sufficient to establish the identity of a literary work.[1] Key to Savile's account of the story is the claim that Menard has composed a story with a different meaning than that of the story composed by Cervantes and that, therefore, although the texts are notationally identical, Menard has produced a different work of literature (Savile 1971: 21-22).

Goodman has a brief and fairly obvious defense to the thought underlying Savile's contention there are two works here. That a single text can have more than one interpretation is, of course, no evidence for claiming there is more than one literary work at hand. Indeed, Goodman and Catherine Elgin offered precisely this defense in 1988. They observe that, even if we construe Menard as having the ambition to create a text notationally identical with Cervantes' and even if that ambition were achievable, what we should then conclude is that

> what Menard wrote is simply another inscription of the text. Indeed […] if infinitely many monkeys […] produce[d] a replica of the text […] that replica […] would be as much an instance of the work Don Quixote as Cervantes' manuscript, Menard's manuscript, and each copy of the book that has ever been or will be printed. (Elgin – Goodman 1988: 62)

A second attempt to use the story to challenge Goodman is found in Arthur Danto's 'Artworks and Real Things' (Danto 1973) to assist him in motivating the problem that was to become the cornerstone of his later book, The Transfiguration of the Commonplace (Danto 1981). The book's title suggests the contents of both the article and the book well enough: in them Danto asks us to consider how it can come about that commonplace objects become works of art; and he provides an answer, namely, that commonplace objects become works of art, if they do, when they are transfigured by being taken as being 'about something' or, in other words, by being 'interpreted' in terms of a theory of art.

Of course, the very idea that commonplace objects can just become works of art, in any manner or by any means at all,

---

[1] There are two previous sightings of the Borges story. Zemach 1968 discusses the story in connection with offering an account of time sensitive properties of works of art. Agassi 1970 also mentions the story, but in a quite different context with quite different implications. Thanks to Gary Iseminger for pointing me to this latter sighting.

needs some motivating. It is not enough that Duchamp puts a bottle rack forward in a gallery or museum, even with a title attached. Nor will it suffice that many members of the museum going public, gallery owners, art-lovers, and purchasers of art acclaim the bottle rack as an 'artwork'. For, as B. R. Tilghman (1982: 293) puts it, "there is no demand grounded either in logic or in sensibility that obliges us to the same". If philosophy is to settle this – and I am not sure it can – what is needed is an argument or a compelling and incontrovertible case, something that does appeal to logic or analysis of sensibility. Duchamp's 'Readymades' do not fit the bill precisely because they are the controversial objects that arguments or appeals to cases are supposed to help us come to resolve.

Rather than a compelling case, Danto offers an argument. He reasons that, if such objects could become works of art, the change in status could not be in virtue of any changes in their perceptible properties (for the simple reason that there are no such changes). Moreover, whatever could bring about the change in status, it must be something that is involved, essentially, in anything's being a work of art. That is, Danto reasons that the identification of something - when it is a work of art- as the work of art it is, cannot be the result of an inference from a description of only its perceptible properties. Correlatively, Danto (1973: 5-7) thinks it must be possible for there to be different works of art that are perceptually indiscernible from each other.

What Danto is thought to have done here is show that identification of works of art demands specification of not only intrinsic but also extrinsic, specifically relational, properties of the work. This view may be true. But Danto has not shown it by his use of Borges' story. As Tilghman points out, "the only ground," Tilghman (1982: 298) writes, "for saying that Borges has contributed to literary ontology by discovering that [historical and critical matters bearing on interpretation] are constitutive [of the identity of a work] is the claim that Menard has really written a new work." But, the grounds Danto offers for the latter claim just are the different relations between Menard and Cervantes and their two (albeit notationally identical) texts, the relations between the texts and what would be their place in literary history were they different works, and so on. In fact, as Tilghman shows, Danto's connection between

constitution and individuation appears to rest on a circular argument or none at all, "to come as a package".

David Lewis (1978) provides a third route to challenging the Goodman position. Lewis maintains that Borges' story illustrates that where there are "different acts of storytelling" there are "different fictions" because a fiction is "a story told by a storyteller on a particular occasion" (Lewis 1978: 39). However, as George Bailey has observed, what bearing Lewis's claims have on issues about the ontology of artworks depends on whether we construe the expression 'different fictions' to mean 'different works'. If not, then the fact - if it is one - that Menard has written a different fiction from the fiction written (on another occasion) by Cervantes in no way counts against there being just one (version of) the Quixote. But, if 'different fictions' does mean 'different works', Lewis is faced with the serious difficulty of denying, implausibly, that literary works can have multiples. Lewis, Bailey (1990: 349) notes, "tries to avoid this by noting that producing a copy is not an act of storytelling […] [but then, and even more implausibly] by his definition copies are not instances of fictions."[2]

2. MENARD'S INTENTIONS AS UNDERSTOOD IN THE ONTOLOGY DEBATE

Although the discussion so far seems to favor Goodman's stance in the ontology debate, that is far from my purpose in this essay. We can take a step towards the actual purpose by noting that, in the middle of this debate about work identity in art has been an assumption about the nature of Menard's intentions in producing the chapters of the Quixote that he manages.

In Flint Schier's discussion of the story we find a rehearsal of the claims first stated by Savile and later implied by Danto. But what he adds is revealing: he characterizes Menard's intention as that of "creat[ing] a work that will be the same, word for word, as Cervantes' Don Quixote." And, he concludes, "If Menard had succeeded in his mad ambition…his work would not be Don Quixote by Cervantes, but rather the emanation of a peculiar symbolist poet of the early twentieth century" (Schier

---

1986: 28). Pretty obviously, Schier's conclusion is offered as a criticism of Goodman, suggesting that Menard has, indeed, produced a notationally identical text that is a new work of literature. But the importance of Schier's contribution lies elsewhere; for he is the first in the debate to try to articulate the content of Menard's intention. This characterization has become, with few exceptions, the standard story about Menard's intentions among parties to the debate.[3] If one thinks Menard realized his intentions, so described, one sides against a generally Goodmanian ontology; if one thinks Menard had not realized his intentions, so described, one supports a roughly Goodmanian line on ontology.

Christopher Janaway (1992) offers an enriched account of Menard's intentions and examines how doing so helps sort things out ontologically.[4] Janaway asks us to imagine the case of a psychological experiment in which two people – let us call them A and B – are given slips of paper. A is told to write down the name of a color or arrange a set of words into a string and B is told to write down what she thinks A will write down (Janaway 1992: 75). What is there to count against our saying that both A and B are composing even if, on the bizarre chance it should happen that what is written down on the two slips of paper coincides exactly?

We might be inclined to say this is a case in which either both or neither is composing. That is, if writing down a color word or arranging a set of words into a string counts as composition at all (or not), it should count (or not) both the case in which there is no guiding idea and in the case in which there is a guiding idea, namely that of writing down what the other person will write down. But in fact the latter seems to have greater claim to be a case of composition than the former.

Neither of these descriptions of Menard's intentions squares easily with Borges' story. (I will argue later that they are also inadequate as accounts of authorial intention, considered more generally.) In this regard notice first how Borges sets out Menard's project: "[Menard] did not want to compose another Quixote – which is easy – but the Quixote itself". Nor did he intend to copy it out; "his admirable intention was to produce a

---

[3] Even Elgin and Goodman accept it. They seek only to separate the question of Menard's achievement from the question of his intention (Elgin – Goodman 1988: 61).

[4] His essay is a reply to Wreen 1990.

few pages which would, word for word and line for line, coincide those of Miguel de Cervantes" (cf. Borges 1964: 39). Now notice that there are two parts to this intention. Menard intends a particular product, namely, that there should be some pages that coincide with those of Don Quixote. But Menard also intends to compose the Quixote: on the one hand this was not to be a copy, not to be plagiarized; but, on the other, the text was not to coincide entirely accidentally with the Quixote. The product was intended to be the result of the intended activity.

Schier's account of Menard's intentions clearly does not measure up, at least when taken as an account of Menard's intentions consistent with Borges' story. This is both because it fails to conform to the to the first part of Borges' characterization (in that it makes Menard aim at more than that at which Borges has Menard aim) and because it disregards the second part, the intention to compose, altogether.

Janaway's account of Menard's intentions might be thought to fare better because it appears to give us a handle on why what Menard is doing could be described as 'composing'. But, on reflection, this will not do. Something critically important is missing from Janaway's case. As Borges tells it, the idea is not that Menard (1) intends to write down some words and (2) hopes the resulting text coincides with some stretches of Don Quixote. One way to characterize what Janaway has omitted is this: there must be an intended link between the writing down and the nature of the text produced in order for us to have a case of 'composing'. This is part of what I meant when I earlier wrote that the product is intended to be the result of the intended activity. Another way of putting this point [5] is this: Menard may have peculiar literary ambitions, and it may be that some of his literary ambitions are quixotic; but they are literary ambitions. The conjunction of (1) and (2) do not express a literary ambition. At best they can be used to set up a 'thought experiment' of a kind that is a staple of ontological discussion.

In contrast, Borges' statement of Menard's ambition – "to compose the Quixote" – seems peculiarly telling. Yet, what can he have meant?

---

[5] At any rate, I think these had better be close to the same point. More precisely we may expect that an analysis of what it is to compose a literary work and an analysis of what it is to put literary ambitions into motion will turn out to be the same analysis.

3. AUTHORIAL INTENTION AS IMPLIED IN BORGES' STORY
3.1. CONSTRAINTS ON ACCOUNTS OF AUTHORIAL INTENTIONS
Before trying to formulate Menard's intentions more explicitly, I believe it would do well to consider a pair of constraints on any formulation we might provide.

The first constraint I have in mind is this. There is a sense in which not even Cervantes can have intended to compose Don Quixote. And whatever constrains Cervantes in this respect should also constrain Menard, so long as what we imagine is that he intends to compose the Quixote. Now I do not see why we cannot imagine that he intends something, analogous to what Cervantes intended, that results (however improbably) in producing the Quixote. For I do not think we know enough about the world to predict the Quixote could just never appear in this way; nor do I think we know what we would actually say about it if it did (despite some glib claims by certain philosophers in an enthusiasm of anti-Goodmanian fervor). And so I think if Menard intends to produce the Quixote, under this constraint on what that can mean, well and good.

In this regard, it might be helpful to think of composing as analogous to solving a problem in algebra. One cannot set out to solve a particular algebra problem and, simultaneously, intend that the expression of the solution just be '$x-1$'. One can, of course, intend to express the solution to the problem as '$x-1$'. But then one will have already worked out the solution and cannot be setting out to solve it. Composing a novel, poem, and so on, is like this: when setting out to compose one does not know how the composition will turn out. Borges is clearly in touch with this idea, for he has Menard write the following in a letter:

> When I was ten or twelve years old, I read it, perhaps in its entirety. Later, I have reread closely certain chapters, those which I shall not attempt for the time being. I have also gone through the interludes, the plays, the Galatea, the exemplary novels, the undoubtedly laborious tribulations of Persiles and Segismunda and the Viaje del Parnaso [...]. My general recollection of the Quixote, simplified by forgetfulness and indifference, can well equal the imprecise and prior image of a book not yet written. (cf. Borges 1964: 41)

The algebra analogy can also serve to remind us that, just as actually solving a problem in algebra does not happen by

accident but as a result of the intention to solve a problem, composing the Quixote, for example, does not happen by accident but as the result of the intentions of the author, however we end up characterizing them.

Borges suggests a second constraint on any adequate formulation of authorial intention when he writes that Menard considered two ways he might arrive at composing the Quixote; the first, dismissed by Menard as "too easy", would be to become Miguel de Cervantes; the second would be to compose the Quixote out of his own (20th Century) experience. Any way we have of coming to terms with either of these compositional routes will have to deal with the issues:

A. what materials of writerly composition were available to Menard or Cervantes;
B. what aims could have been in view for Menard or Cervantes;
C. whether the materials deployed are appropriate to the aims in whose service they were deployed.

This third issue initially may seem more a matter of critical judgment rather than any part of the content of literary ambition. So set it aside for the moment.

The first two issues appear to be subjects for inquiry in literary history. Clearly Menard could have Cervantes' material available. In fact Borges takes pains to stress what Menard did to acquire them. Menard might even be able to choose to adopt Cervantes' aims, but perhaps only nostalgically. This is why, even if it is nonsensical on other grounds, Menard should think of the alternative of becoming Cervantes: were he to do so, his aims would just be those of Cervantes, and writing the Quixote with Cervantes' aims would be considerably less challenging.[6] Alternatively, Menard might attempt to put Cervantes' materials to use to other, 20th Century, ends. I suppose it must be something like this that Menard chooses to do. But, characterized in this way, there is no particular philosophical oddity in Menard's project. It is quixotic, of course; and what makes the project quixotic is that Menard chooses to use exactly the same materials, in exactly the same ways, for these

---

[6] Not without challenges, however, as we shall see in the next two sections. One way to put the matter here would be just to say that it is unclear whether even Cervantes could have had the aim of composing Don Quixote, even though he did compose it and even though his composing it was no accident.

other ends. I will soon show that there is a way to account for our sense Menard's intention is more than merely a little odd, that there is in fact something inconsistent or incoherent in the offing, by appeal to the third issue described above. For, I believe, we sometimes have to attribute to authors, as part of their normal intentions, a specific intention with respect to the appropriateness of materials to aims.

3.2. AN ABSTRACT FORMULATION OF AUTHORIAL INTENTIONS

Within these constraints I now suggest the following schematic of any authorial intention:

> (a) I intend to assemble (literary) materials to realize some project;
> (b) I intend that the assembled materials will allow, induce or invite my audience, the readers, to have certain reactions or reaction-types; and
> (c) I intend that the readers' reactions come about as a result of the fact that the materials actually do fit my aims and [some version of a judgment of fitness claim] and [possibly, some version of an adaptivity claim].

Some general remarks are in order. Clearly this formulation allows us to get at two of the themes I claim one has to contend with in order to understand what Menard is supposed to have intended: (one) what materials of writerly composition were available to Menard or Cervantes, and (two) what aims could have been in view for Menard or Cervantes.

I have introduced the third clause for two reasons. First, I think an author would think she had failed to realize her ambitions were she to discover that readers had reacted exactly as she intended but demonstrably had done so because of other features of the composition than those she saw as connected to those reactions, or that they had reacted as she intended but for reasons that had nothing whatever to do with the composition itself. Secondly, the third clause constitutes a relatively uncontroversial expression of the third theme mentioned above as possibly necessary for understanding Menard's and Cervantes' intentions: (three) whether the materials deployed are appropriate to the aims in whose service they were deployed.

The second and third conjuncts of the third clause are meant to be placeholders for substantive beliefs that render robust even this abstract formulation of authorial intentions. Without beliefs of the kind that would fill in these gaps, we might well

wonder if the third clause accurately captures all possible authorial ambitions with respect to the products of their activities. Noting two things can bring out what I have in mind.

First, in addition to intending that readers' reactions come about as a result of the fact that the materials actually do fit the author's aims, an author might also intend that this come about as a result, further, that the readers have understood – by exercising their own critical judgment – that the various conventions, grammatical devices and other writerly strategies that are the author's 'materials' do fit the author's aims. In contrast, another author may wish the readers' reactions to come about as a result of the fact her materials fit her aims, but not intend the readers be aware of how that happens, that it not happen as a result of the readers exercising critical judgment. Such beliefs as these are what I have called 'judgment of fitness claims'.

Second, authors may hold beliefs, just as anyone else might, concerning the degree to which adaptivity of literary materials to literary aims is historically conditioned. Such beliefs may be weaker or stronger. A relatively weaker and uncontroversial belief of this kind is the view that artistic aims not possible in one time period may become possible at a later time. This kind of belief is what fills in the space for what I have called 'adaptivity claims'. Such a belief can become part of what gives full expression to the content of an author's intentions in the sense that, without reference to such a belief we may not be able to characterize adequately and fully the author's intentions.

I have posed these last points as a pair of options with respect to which an author might adopt a point of view with regard to qualifying the intention that readers' reactions come about as a result of the fact her materials do fit her aims. One might think we should settle philosophically which ways of responding to each of these options are correct. I do not think this is so and, as I will now show, it is for that reason I do not think we can determine philosophically what makes Menard's intentions incoherent. What I do think philosophy can do for us here is make clearer what are the elements of authorial intention, what elements must be present and, where options are available to an author, what those options consist in and severally entail.

Suppose we adopt a fairly strong formulation of the judgment of fitness claim in order to give an account of literary ambitions. A strong formulation would hold that, to be described

as having literary ambitions, an author must be thought of as intending to assemble materials in such a way that allows, indeed invites, one's audience not only to have certain reactions and reaction-types, as a result of the fact the materials do fit the aims, but also to exercise its own critical judgment about what materials are appropriate to what aims.[7] If we adopt this strong formulation, we may be led to say that one cannot reproduce – let alone set out to reproduce – an already existing work unless one has exactly the same aims as the author of that existing work. We certainly can be led to this conclusion if we also adopt a strong view about the adaptivity of materials and ends. If we – in contrast to the author – believe only one precise set of means is adaptable to any precise set of aims, then we are committed to the view that no author could attempt to invite her readership to do what she invites it to do by assembling for new aims the same materials already deployed in one work for another set of aims. She would have to suppose them incapable of seeing that the means are fit for earlier aims while supposing them capable of seeing, and indeed inviting them to see, that the means she deploys are fit for her new aims.[8] On these grounds there seems to be real evidence of something inconsistent or incoherent in the Menard's project (but that inconsistency or incoherence has yet to be spelled out precisely).

To arrive at this point we have adopted a strong formulation of the judgment of fitness claim as part of the expression of normal authorial intentions and adopted a strong position on the degree to which adaptivity of materials to aims is historically conditioned. A weaker version of the judgment of fitness claim could hold that the author intend only to exercise her own judgment concerning fitness of materials to aims without expecting the same from her readers. That is, this author could be described as intending to assemble materials in such a way that allows and invites her audience only to get the point (because the materials do fit the aims) without consciously deciding for themselves that the materials fit the author's aims. This will still render Menard-like aims as quixotic if we still

---

[7] Here I am developing an idea about authorial intentions regarding 'fit' (as between materials and aims) to be found in Wollheim 1984.

[8] Indeed, this is Wollheim's way of handling the matter. One also should note our reliance on the view that, to intend to do a one cannot know one cannot do a. This is likely to be the least controversial part of Wollheim's way of handling the matter.

accept the strong adaptivity claim. But if the author does not likewise adopt the strong adaptivity claim,[9] there will be nothing incoherent about her intentions. And if we drop the strong adaptivity claim, not only are her intentions coherent, they also lose even the appearance of being quixotic. This suggests a line we can take in capturing the incoherency in Menard's intentions.

3.3. WHY (AND HOW) MENARD'S INTENTIONS ARE INCOHERENT

We are able to demonstrate Menard's intentions are incoherent in part by attributing to him a strong version of the judgment of fitness claim (noted in italics). Menard, we may say, intends (in part) the following:

> I intend that the readers' reactions come about as a result of the fact that the materials actually do fit my aims AND that they have this reaction as a result of exercising their own critical judgment that the materials are appropriate to the aims.

Then, to complete our account of what makes his intentions incoherent, we attribute to him a strong version of the adaptivity claim. Menard, we may say, believes the following:

> Only one precise set of means is adaptable to any precise set of aims.

Now, one might think we want another way to arrive at 'Menardian incoherency,' if I may call it that, that does not depend on his holding any version of the adaptivity claim. At least we may think that it shouldn't fall out that Menard's intention is incoherent only in case the intending agent is committed to the strong adaptivity claim.

This is not because the strong version of the adaptivity claim is false. It is false. If the notion of 'literary means' has scope to include, for example, the specific forms of direct authorial address used by both Henry Fielding and Jorge Amado, whose uses of this same technique are to very different effects, then the strong version of the adaptivity claim has to be false.[10]

---

9 The situation may be different if she merely has doubts about whether she can use the old materials for new aims. For a discussion of this issue, and related matters about 'side effects' not directly addressed in the present essay, cf. Mele 1992.

10 Of course, as we have noted, there can be weaker versions of the claim. It might be true that some fairly specific literary techniques are adaptable to only a re-

The reason we may want Menard's intention to come out incoherent even if Menard believed no version of the adaptivity claim, true or false, is that otherwise, Menard's incoherence is contingent on the merely accidental fact the intending agent happens to believe something about the adaptivity of literary means to literary (and other) ends. And this may seem extraneous. But, if this is what we want, I believe we are mistaken. Look again at the passage in which Menard writes about how he conceives of the project:

> Once that image [of a book not yet written] (which no one can legitimately deny me) is postulated, it is certain that my problem is a good bit more difficult than Cervantes' was. My obliging predecessor did not refuse the collaboration of chance: he composed his immortal work somewhat à la diable, carried along by the inertias of language and invention. I have taken on the mysterious duty of reconstructing literally his spontaneous work. (Borges 1964: 41)

Here Menard points out three "obstacles" to his project. The first is that he is allowed to make variations "of a formal or psychological type". The second is that he is "oblige[d]..to sacrifice these variations to the 'original' text and to reason out this annihilation in an irrefutable manner". To these "artificial hindrances", Menard writes, there is added a final obstacle, "of a congenital kind":

> To compose the Quixote at the beginning of the seventeenth century was a reasonable undertaking […] at the beginning of the twentieth, it is almost impossible. (ibid.)

To be sure, this is not the strong adaptivity position we have been working with to arrive at our account of Menardian incoherence; and so our account must be adjusted. But it is clear from the passage that Menard holds some version of the adaptivity claim. I think we can capture the version he holds as follows:

> Only one precise set of means is reasonably adaptive to any precise set of aims.

stricted range of effects. I do not dispute that. There is room for a lot of subtlety here.

It is because Menard believes this (also false) version of the adaptivity claim that he conceives his entire project with the intention to undertake, as though it were reasonable, something which also he believes in fact to be unreasonable. And this, at least, is a kind of incoherence.

That he has this incoherent intention does not, however, make the appearance of a text, produced in this way, impossible. For, as I have noted above, I do not think we know enough about the world to predict Quixote, or two chapters and part of a third of it at any rate, could just never appear in this way or something like it. Nor do I know what we would/should actually say about it if it did.

So, I conclude, we can extract an interesting account of authorial intentions from Borges' story and, by so doing, we can show what is quixotic about Menard's intentions. Because the incoherency of his intentions turns on empirical facts about what Menard believes, or could believe, we cannot show his intentions must be incoherent, as a matter of logical or conceptual necessity. This fact is part of what makes visible the possibility that both Borges' implied account of authorial intention and his description of one way in which authorial intentions can misfire – both prominent features of the story – are obscured for us by using the story to motivate a debate in the ontology of art.

4. AN OBJECTION: THE APPARENT NEED TO RESOLVE THE ONTOLOGY DEBATE

A plausible objection to the line I have been taking is that my argument succeeds only because I have been overly vague about what counts as 'materials' and 'aims'.[11] The objection might proceed by suggesting that the best way to make these notions precise is to state identity criteria for materials and then ask whether aims supervene on materials or materials underdetermine any aims realizable.

[11] Peter Lamarque suggested this line of objection to me in comments he made on an earlier version of this paper at the American Society for Aesthetics, Pacific Division, in March 1999. I am grateful to him for spelling this out so clearly and allowing me to read his comments in detail. Readers cognizant of recent discussions of actual-author intentions versus hypothetical-author intentions in the context of settling questions of interpretive scope will find the terms of this objection familiar. See, for example, Carroll 1992; Carroll 2000; Iseminger 1996; Levinson 1992; Stecker 1993; Stecker 1997; Trivedi 2001.

A favored way of talking about literary works of art would then lead us to decide whether by 'materials' we mean inscription- or sentence-types, sentence meaning, or utterance meanings.

If the materials Menard used are identified as inscription- or sentence-types, then Menard and Cervantes use the same materials and, clearly, these materials underdetermine aims. Sentence-types can bear multiple meanings and not every utterance of a sentence-type realizes the same communicative aim.

This suggests we include sentence meaning in the identity conditions of materials. But sentence meaning is determined by the semantics of a language so we can, once again, have different realizable aims. Here the objector can supply familiar cases involving indexicals to show that sentence meaning underdetermines the aims speakers may have.

Suppose we identify materials with utterance meaning. Now aims do look like they supervene on materials, for any change of aim would imply a change of materials.

None of these seem to be what Menard has in mind. Menard's agonizing and endless drafts are not consistent with identifying his materials with inscription- or sentence-types. Menard's explicitly stated intention to compose the Quixote is not consistent with identifying his materials as sentence meanings. Nor finally, will it do to identify his materials with utterance meanings: this latter would require Menard to become Cervantes, which, you remember, Menard dismisses as "too easy"! The objector can now add that Borges' narrator points out that the final product of Menard's efforts is not identical to the original; it is, as he puts it, "infinitely richer", "more ambiguous", it is in a different style, it makes different claims about history, and so forth.

So, it may be said, there is a way to show Menard's intention is incoherent of necessity. This would then justify us in looking at his achievement, whatever it is he produces, free from any constraints having to do with his intentions in producing it. His apparent achievement, as described by the narrator, is to have produced something seemingly distinct from the Quixote, with different properties, including a different meaning. The achievement can just about be described as a new work built on the material of an earlier writer.

It is precisely here that we need to invoke a distinction between work and text. For example, one familiar story about this would insist that, if we take the identity conditions of the

material to be text-identity, including not only notational but also semantic identity, then we can describe Menard as reproducing the same textual material as Cervantes. If we take work-identity to be distinct from text-identity and to incorporate features like utterance meaning and circumstances of origin, then Menard can be given credit for producing his own work different from Cervantes. There is nothing more perplexing about this achievement than the fact that two speakers can use the same sentence-types to make distinct statements, with different aims and subject to different interpretations.

To sum up then, the first argument we have been tracking in stating the objection looks like this:

(1) Menard either aims at producing
(A) the same sentence- or inscription-types as Cervantes,
(B) the same sentence meanings as Cervantes or
(C) the same utterance meanings as Cervantes.
(2) The narrator should be taken at his word with respect to what Menard does and says about what he is doing.
(3) Menard agonizes over multiple drafts.
(4) Premise (3) is inconsistent with alternative (A) in (1).
(5) Menard insists he wants to "compose the Quixote".
(6) Premise (5) is inconsistent with alternative (B) in (1).
(7) Menard dismisses the idea of becoming Cervantes as "too easy" for his purposes.
(8) Premise (7) is inconsistent with alternative (C) in (1).
(9) Insofar as premise (1) exhausts the possibilities with respect to Menard's intentions, the narrator's description of Menard's intentions is incoherent.

Having shown Menard's stated aims are incoherent, we are free to consider independent arguments concerning whether his achievement is coherent. The argument rehearsed above as an example can be presented thus:

(10) Text-identity is determined by not only notational but also semantic identity. Work-identity incorporates features like utterance meaning and circumstances of origin.
(11) The narrator should be taken at his word with respect to what Menard produced.
(12) So, we can gloss what happens in the story this way: Menard reproduces the same textual material as Cervantes but Menard produces his own work different from Cervantes'.
(13) The achievement description in (2) is coherent. Indeed it is the same sort of explanation one might give of some of the

achievements of Shakespeare who famously used and reworked material of his predecessors while still producing original works.

And our objector's conclusion is the conjunction of (9) and (13): Menard's stated aims are incoherent but the narrators' description of Menard's achievement is coherent.

5. A RESPONSE

In assessing these arguments the first thing we might want to notice is that (9) and (13), conjoined with (2) and (11), imply that Borges' narrator is confused. This is surely possible; a fictional narrator might be confused about many things. What we must claim here is that Borges' narrator fails to notice, at least in the case of Menard, that we cannot draw the same connections between intentions and attributions of achievement we ordinarily do. Borges' narrator does think that Menard's intended products are the result of his intended activities.

There is a problem with this. If the interpretive claim is that the narrator is confused in this way, we want more than Borges presents in the story by way of support for the claim. There should be some independent textual evidence of the narrator's disposition to confusions of this kind. Borges, as I have just said does not provide the relevant evidence. Of course, another possibility is that Borges himself is confused. Short of either of these hypotheses, however, we might want to see if more could be done by way of analysis of the story. So, let us hold (2) and (11) fixed.

To clear away one unnecessary source of complexity, let me note that I am not at all concerned with the argument for (13) except insofar as the entire argument is a symptom of things gone seriously awry. True enough, the objector's conclusion, as I have presented it, depends on the strength of the inference from (10) through (12) to (13). But (10) and (12) are subject to considerable dispute among philosophers working on the ontology of works of art and Menard's 'achievement' has been described variously as plagiarism (Tilghman 1982; Bailey 1990), copying (Bailey 1990; Sparshot 1982: 179; Wreen 1990), and the coincidental writing of the same work (Janaway 1992; Currie 1989: 120-124). I take no position on this matter.

What I do want to take up is the fact that, as the objection is worked out, we are entitled to undertake an argument concerning whether Menard has or has not produced a new

work of literary art only because we have concluded that his intention is incoherent and therefore we may ask about his achievement without reference to his intentions or ambitions. But I have insisted that, in any robust notion of authorial intention, we must understand the products of those intentions as having come about non-accidentally as results of activities undertaken under the direction of the author's intentions. If I am right we should be able to detect something fundamentally wrong with the argument from (1) through (8) to (9).

And indeed we can. Here, again, is that argument:

(1) Menard either aims at producing
(A) the same sentence- or inscription-types as Cervantes,
(B) the same sentence meanings as Cervantes or
(C) the same utterance meanings as Cervantes.
(2) The narrator should be taken at his word with respect to what Menard does and says about what he is doing.
(3) Menard agonizes over multiple drafts.
(4) Premise (3) is inconsistent with alternative (A) in (1).
(5) Menard insists he wants to "compose the Quixote."
(6) Premise (5) is inconsistent with alternative (B) in (1).
(7) Menard dismisses the idea of becoming Cervantes as "too easy" for his purposes.
(8) Premise (7) is inconsistent with alternative (C) in (1).
(9) Insofar as premise (1) exhausts the possibilities with respect to Menard's intentions, the narrator's description of Menard's intentions is incoherent.

Note first that premise (2) is what underwrites our confidence in premises (3), (5), and (7). These latter are what are appealed to in determining that none of the ways of specifying Menard's materials listed in premise (1) can be sustained. And that , in turn, is the substance of the consequent of the conclusion, (9).

This argument is, however, only as strong as our confidence in the antecedent of the conditional that forms the conclusion. It might be more perspicuous to note that this antecedent functions as an assumed premise in the argument. But now, if we think it implausible that Borges' narrator (or Borges himself) is confused, we should conclude that premise (1) does not exhaust the possibilities with respect to Menard's materials.

This is a conclusion that I think we should also arrive at quite naturally from considerations of what authors themselves think of as their "materials." The thought behind the objection is right, that this term captures many different kinds of things. The term

is vague. But the correct way to make it more precise is to offer detailed examples.

Borges had, I believe, already done this in his story. It may be sufficient – both to show how we should make the concept of writerly "materials" more precise and to demonstrate what must get obscured if we focus on ontological questions about works of art – to list the kinds of things Menard took to be his (and Cervantes') materials. Remember what Menard says he did to acquire an "equal [to] the imprecise and prior image of a book not yet written":

> I have reread closely certain chapters, those which I shall not attempt for the time being. I have also gone through the interludes, the plays, the Galatea, the exemplary novels, the undoubtedly laborious tribulations of Persiles and Segismunda and the Viaje del Parnaso.[12]

Among the materials Menard consults then, the other works of Cervantes, are where he will find the multiplicity of sentence structures, the techniques for achieving depth of character, the simplifications of action, and the commitment to detailed descriptions (achievable in prose but not in poetry – at which Cervantes was apparently not very good) that characterize the writing of Cervantes. Then too, there is a rhythm and sound in Cervantes' prose that would be the appropriate object of Menard's attention. These elements, and things like them, are what we need to think about when we consider what are an author's "materials."

But these elements, and things like them, disappear from view when we hold that the materials an author deploys are inscription- or sentence-types, sentence meanings, or utterance meanings. Under the influence of a conception of authorial intention that is friendly to the ontology debate, a more robust and adequate conception of authorial intentions – where the articulation of those intentions makes appeal to such concepts as 'materials', 'aims', and 'fit' – becomes unavailable to us.

---

[12] I might also note that, even here, Borges is having us on. La Galatea and the two surviving plays antedate the Quixote. But the Exemplary Novels and the Viaje del Parnaso both come some ten or more years after Cervantes wrote Part I of Don Quixote. And the Travails of Persiles and Seigismunda was written a year after Part II of the Quixote.

BIBLIOGRAPHY

Agassi, Joseph. 1970. 'Philosophy as Literature: The Case of Borges' in *Mind* 79: 287-293.

Bailey, George. 1990. 'Pierre Menard's Don Quixote' in Iyyun, the *Jerusalem Philosophical Quarterly* 39: 349.

Borges, Jorge L. 1964. 'Pierre Menard, Author of the Quixote' in *Labyrinths: Selected Stories and Other Writings*. New York: New Direction Publishings: 36-44.

Carroll, Noel. 1992. 'Art, Intention, Conversation' in Iseminger, Gary (ed.) *Intention and Interpretation*. Philadelphia: Temple University Press: 97-131.

Carroll, Noel. 2000. 'Interpretation and Intention: The Debate between Hypothetical and Actual Intentionalism' in *Metaphilosophy* 31(1-2): 75-95.

Currie, Gregory. 1989. *An Ontology of Art*. New York: St. Martin's Press.

Danto, Arthur C. 1973. 'Artworks and Real Things' in *Theoria* 39(1): 1-17.

Danto, Arthur C. 1981. *The Transfiguration of the Commonplace*. Cambridge (Mass.): Harvard University Press.

Diamond, Cora. 1988. 'Losing Your Concepts' in *Ethics* 98(2): 255-277.

Elgin, Catherine and Nelson Goodman. 1988. *Reconceptions in Philosophy and Other Arts and Sciences*. London: Routledge.

Goodman, Nelson. 1968. *Languages of Art*. Minneapolis: Bobbs-Merrill Company.

Iseminger, Gary. 1996. 'Actual Intentionalism vs. Hypothetical Intentionalism' in *Journal of Aesthetics and Art Criticism* 54(4): 319-326.

Janaway, Christopher. 1992. 'Borges and Danto: a reply to Michael Wreen' in *British Journal of Aesthetics* 32(1): 72-76.

Levinson, Jerrold. 1992. 'Intention and Interpretation' in Iseminger, Gary (ed.) *Intention and Interpretation*. Philadelphia: Temple University Press: 221-256.

Lewis, David. 1978. 'Truth in Fiction' in *American Philosophical Quarterly* 15(1): 37-46.

Mele, Alfred. 1992. 'Recent Work on Intentional Action' in *American Philosophical Quarterly* 29(3): 199-217.

Savile, Anthony. 1971. 'Nelson Goodman's Languages of Art: A Study' in *British Journal of Aesthetics* 11(1): 3-27.

Schier, Flint. 1986. *Deeper into Pictures*. Cambridge: Cambridge University Press.

Sparshott, Francis. 1982. *The Theory of the Arts*. Princeton: Princeton University Press.

Stecker, Robert. 1993. 'The Role of Intention and Convention in Interpreting Artworks' in *Southern Journal of Philosophy* 31(4): 471-489.

Stecker, Robert. 1997. 'Hypothetical Intentions and Implied Authors' in Stecker, Robert *Artworks*. Philadelphia: Pennsylvania State University Press: 186-212.

Tilghman, B.R. 1982. 'Danto and the Ontology of Literature' in *Journal of Aesthetics and Art Criticism* 40(3): 293-299.

Trivedi, Saam. 2001. 'An Epistemic Dilemma for Actual Intentionalism' in *The British Journal of Aesthetics* 41(2): 192-206.

Wollheim, Richard. 1984. 'On the question "Why painting is an art?"' in *Aesthetics: Proceedings of the 8th International Wittgenstein Symposium*, Part I. Vienna: Hölder-Pichler-Tempsky: 101-106.

Wreen, Michael. 1990. 'Once is Not Enough?' in *British Journal of Aesthetics* 30(1): 149-158

Zemach, Eddy. 1967. 'A Stitch in Time' in *The Journal of Value Inquiry* 1: 223-241.

ALBERTO VOLTOLINI

A Syncretistic Ontology of Fictional Beings

Abstract: : In this paper I will sketch a syncretistic theory of fictional beings, according to which fictional entities are compounds both of a make-believe process type and of a set of the properties corresponding to those mobilized in that process. This theory is syncretistic for it tries to combine the virtue of the two main contemporary approaches to *ficta*, the neo-Meinongian and the artifacualist. But it is also more conciliatory than that for it tries to combine both an antirealist and a realist perspective on such entities.
Keywords: ontology of art; works of art; fictionality.

In the camp of the believers in fictional entities, two main paradigms nowadays face each other: the neo-Meinongian and the artifactualist.[1] Both parties agree on the idea that ficta are abstract entities, i.e. things that exist (at least in the actual world) even though in a non-spatiotemporal way. Yet according to the former paradigm, ficta are entities of a Platonic sort: either sets of properties (or at least 'one-one' correlates of such sets) or generic objects. According to the latter paradigm, fictional beings are instead abstract artifacts, in the sense that they are cultural constructions like games, laws and institutions.

Traditionally, these paradigms are conceived by their proponents as mutually exclusive. In what follows, however, I will try to show that this conception is ungrounded. For a fictional entity is a compound entity made both of a property set and of the cultural practice-type that makes its own existence possible. This makes a fictum at least a 'many-one' correlate of a set, insofar as different practice-types may turn the same set of properties into different fictional individuals. In this sense, the

---

[1] There is at least another paradigm, namely possibilism. Yet there are serious drawbacks in conceiving fictional entities as possible entities. On this point, cf. Thomasson (1999: 17-18) and Voltolini (1994).

present proposal is ontologically syncretistic, for it attempts at combining the neo-Meinongian and the artifactualist paradigm.

Yet it is even more conciliatory than that. Recent disbelievers in ficta have maintained that as far as fiction is concerned, there is nothing more than fictional discourse itself, which consists in nothing but make-believe linguistic acts in which we pretend that there are things like fictional beings. Yet I take this make-believe practice precisely as the cultural practice such that a fictum not only depends on it but also is partially constituted by it.

1. Let me start by evaluating the neo-Meinongian position. According to neo-Meinongians, ficta are abstract entities, namely entities that exist albeit in a non-spatiotemporal way (i.e. subsist, as Meinong would have said).[2] Moreover, they are individuated in terms of their properties. Indeed, neo-Meinongians defend a principle of object-generation that appeals to properties themselves: for any collection of properties, there is an object that has all and only them.[3] Such an individuation may be either direct or indirect. In the former case, a fictum is a certain set of properties, or at least a 'one-one' correlate of such a set: for each set of properties, there is just one fictum that is either identical with or corresponds to it, and vice versa.[4] In the latter case, a fictum is not a set, but a generic object, like the Beautiful and the Triangle. Yet properties count also in the individuation of one such object. For a generic object is what possesses

---

[2] For Meinong's thesis to the effect that abstract entities subsist, cf. e.g. Meinong (1960: 79-80). To be sure, in this conception neo-Meinongians part company with Meinong. For Meinong is ordinarily taken to hold that ficta are a subset not of abstract but rather of außerseiende beings, i.e. entities that are existence-indifferent ("beyond being and not-being").

[3] In Meinong this generation principle is expressed as follows: an object corresponds to every being-so (Meinong 1916: 282). For more technical formulations cf. Parsons (1980: 19) and Zalta (1983: 12).

[4] The set-theoretical conception is defended in Parsons (1980). To be sure, Parsons officially defends the possibility of conceiving a fictum as a 'one-one' set-correlate, rather than as a mere set. Cf. Parsons (1980: 18 and fn.1, 54-55). This possibility is thoroughly developed in Castañeda's theory of guises. In Castañeda's theory, individual guises are the ontological results of the application of a certain operator c to a certain set of properties (the so-called guise core). Cf. Castañeda (1989b: 240). So, a guise precisely is a 'one-one' set-correlate. Of course, in Castañeda's picture a fictum is not exactly a guise, but rather a system of guises tied together by a specific type of predication, consociation. Cf. Castañeda (1989a).

in an internal way – it encodes, as some would say – a certain number of properties; in the case of a fictional being, these are the properties which according to a certain body of literature it possesses in an external way – it exemplifies (cf. Zalta 1983: 91-99). In both cases, a fictum is a Platonic entity. This is evident in the latter case – a generic object is something like a Platonic kind – but it also holds in the former case: qua mathematical entity, a property set belongs to the realm of Platonic beings. Thus, according to neo-Meinongians ficta basically are Platonically abstract entities: that is, they are atemporal beings – in all the possible worlds which contain them, they exist non-spatiotemporally – as well as necessary beings – i.e., they subsist in all possible worlds.

According to neo-Meinongians, therefore, property identity is a necessary, and (possibly) also a sufficient, condition for the identity of a fictional being.[5] Yet as I see it, this position is threatened by at least two main problems affecting the thesis that properties provide a sufficient condition for a fictum.

The first problem is the 'no ficta' problem. In actual fact, in individuating either a set or a Platonic kind, properties individuate what neo-Meinongians would call a Meinongian object. Such an object is an entity characterized precisely in terms of the properties that qualify its nature. Let me adopt the convention of using capital letters to speak of such objects. We thus have that the golden mountain is something that is a mountain and golden, the round square is something that is round and square. Now the question is, what makes a Meinongian object a fictional entity? Take the following example (which I borrow from Kripke 1973). Erroneously, former interpreters of the Bible have taken the name 'Moloch' as referring to a mythical monster where, as modern philology has shown, it is actually used in the Bible as a common noun either for kings or for human sacrifices. Now, the philological error notwithstanding, undoubtedly there is a set $M = \{F,G,H...\}$ constituted by the properties F,G,H... former interpreters have mistakenly understood the Bible as assigning to a certain character. Correspondingly, there is a Platonic kind that internally has only those very same properties. Yet neither that set nor this kind is by itself identical with a fic-

---

[5] I say 'possibly' for if a fictum is taken as a 'one-one' set-correlate rather that a set, then of course properties do not yield a sufficient condition for a fictum.

tional character, insofar as within the realm of fictional characters there is no such a thing as Moloch.

A defender of a neo-Meinongian theory of fictional entities might try to circumvent the 'no ficta' problem by saying that a fictum is a Meinongian object whose constitutive properties are the properties effectively mobilized in some body of fiction (a text, a myth). That is, if some body of fiction effectively narrates that an individual possesses certain properties, then either the set containing all those properties or the corresponding kind coincides with a fictional entity. As in the 'Moloch' case the Bible effectively performs no such narration, there is no fictional character as Moloch. Alternatively, the neo-Meinongian might retreat to 'one-one' set-correlates by saying that even if the 'Moloch' case shows that a set is not the same as a fictum, the latter entity can still be identified with a 'one-one' set-correlate.

Yet these replies do not manage to face another problem. The Meinongian abstractionist says that the properties that turn either a set of properties or a Platonic kind internally having only those very properties into a fictional object are the properties mobilized in some body of fiction. Yet the trouble is that there may well be either one and the same set of properties mobilized by some body of fiction or one and the same kind internally having those very properties only and still different characters.

This is the 'many ficta' problem. It is not only the case that either a set of properties or a Platonic kind does not by itself generate any fictum; it may also be the case that either Meinongian entity matches more than one fictum. Take the famous example where Borges imagines that a certain Pierre Menard, totally disconnected from Cervantes, happens to write a text that is word by word identical to Cervantes' Don Quixote. In such a case, one and the same set of properties (or one and the same Platonic kind internally having those very properties only) matches different characters, namely Cervantes' Don Quixote and Pierre Menard's Don Quixote. It would be useless for the neo-Meinongian to appeal to the fact that the properties in question are mobilized in a body of fiction. For one such mobilization yields different characters (Lewis 1978: 39; Thomasson 1999: 56). Moreover, appealing to 'one-one' set-correlates would not do as well, for the Menard case would show again show that one and the same 'one-one' set-correlate can match distinct ficta.

The neo-Meinongian might deny that in such a case two different ficta are in question: insofar as the set of properties / the Platonic kind is the same, so is the fictional object (Parsons 1980: 188). [6] Yet this is hard to swallow. As I said, Pierre Menard is imagined as a subject completely disconnected with Cervantes, namely an individual that in writing his story just happens to repeat the words that were used by Cervantes in writing Don Quixote. Moreover, suppose Borges' case were reformulated in terms of a Twin-Earth experiment. To be sure, notwithstanding their spatial difference Earthians and Twin-Earthians may well share their mathematics. So, the first may well conceive the very same Platonic entities (sets of properties, kinds) that are conceived by the second. Yet it would hard to admit that they share the same fictional characters and more generally the same fictional world (although of course the texts their fictions are made of would be syntactically identical entities).

So, given both the 'no ficta' and the 'many ficta' problem, it turns out that a fictional object cannot coincide with a Meinongian object, whether a set of properties, a 'one-one' set-correlate, or the corresponding Platonic kind.

2. Let us turn then to the other abstractionist position, the artifactualist one. According to this position, though abstract, fictional beings do not belong to any Platonic realm. They live with us, in the same way as games, laws and institutions do. They are indeed cultural abstractions.

From now on, I will focus on Thomasson's position (Thomasson 1999), for it is definitely one of the richest versions of artifactualism (previously Kripke 1973; Van Inwagen 1979; Salmon 1998). As Thomasson has maintained, the cultural nature of fictional beings means that their life is related to the existence of other cultural entities, like agents and their literary works. Indeed, not only they are dependent entities, in the sense that necessarily, if they exist, some other entities exist as well,but also they depend in respectively different ways on agents and their works. This makes it the case that, though abstract, ficta exist in a both modally and temporally different way from Platonic abstracta.

---

[6] In the Menard case, the abstractionist would probably maintain also that there is just one literary work at issue.

To begin with, according to Thomasson the generic notion of dependence just recalled above must be articulated in further subdistinctions. First of all, one must draw a distinction between constant dependence and historical dependence. The first relation is such that one entity requires that the other entity exist at every time at which it exists, while the second relation is such that one entity requires that the other entity exist at some time prior to or coincident with every time at which it exists (Thomasson 1999: 29). Another distinction is that between rigid dependence, that is, dependence on a particular individual, and generic dependence, that is, dependence on something or other of a particular type (Thomasson 1999: 27). On the basis of these distinctions, Thomasson claims that a fictum is an entity that, on the one hand, historically and rigidly depends on the mental act by means of which its creator originally thinks of it, and, on the other hand, constantly and generically depends on some literary work or other in which it is spoken of (Thomasson 1999: 35-36). As a result, in this perspective ficta are neither atemporal nor necessary beings. A fictum comes into existence as soon as a certain creative act of thought generates it, and moreover ceases to exist as soon as there absolutely are no more copies of any literary work in which it was spoken of. Moreover, a fictum exists only in the worlds in which it has been generated by means of that very same creative act ʿThomasson 1999: 7-11, 38).

Armed with these theoretical tools, artifactualists like Thomasson may easily solve both the 'no ficta' and the 'many ficta' problem. As you will recall, the 'no ficta' problem was exemplified by the 'Moloch' case: although a set of 'Moloch-ish' properties is already at one's disposal, there is no such a fictional thing as Moloch insofar as no 'Moloch' story is really recounted in the Bible. Yet it is clear how there might have been such a thing as Moloch: if the Bible had really contained the 'Moloch' story. And this means if at the times in which the Bible was composed there had been someone who had conceived (and accordingly written in the Bible books) Moloch as the protagonist of the story we erroneously think it is actually contained in the Bible. (This is the solution suggested by Kripke 1973.)

The 'many ficta' problem has a similar solution. Remember Pierre Menard's case: two syntactically identical texts however written by two totally disconnected people, Cervantes and Pi-

erre Menard. Such texts imply that there have been two differ-ent mental acts both thinking of a person named 'Don Quixote', one had by Cervantes and another had by Pierre Menard. Given historical rigid dependence of ficta on their creators' mental acts, two different characters are brought into existence, two Don Quixotes: they are distinct ficta even though they are ascribed in such texts all the very same properties (Thomasson 1999: 6-7, 56).

According to this position, properties ascribed to a character in the relevant body of literature do not seem to play any par-ticular role in the individuation of a fictional being. As Thomas-son maintains, a fictum remains the same entity across different possible worlds even if in those worlds its author had decided to ascribe different properties to it, by writing altogether different works (Thomasson 1999:39). For Thomasson a fictum pos-sesses the ascribed properties only in a relative sense: when a sentence involving the ascription of one such property is evalu-ated with respect to a certain fictional context (the relevant body of literature), then it is true that the fictum has it; when the same sentence is evaluated according to a real context, then it is no longer true that the fictum has it (Thomasson 1999: 105-107). Yet for Thomasson, to say that one such sentence is evaluated as true with respect to a fictional context is the same as saying that according to that context, the fictum has the involved property (Thomasson 1999: 107). Thus, she seems to admit that a fictum has in an absolute sense the corresponding relative property – i.e., what could be called an 'in the story'-property. For instance, Othello is not absolutely jealous, he is only jealous with respect to Othello; as a result, it absolutely has the relative property of being jealous according to Othello. Yet given that she seems to acknowledge that a story might remain the same even if it were altered a bit (Thomasson 1999: 109-111), then not only absolute, but also relative properties are contingently possessed by fictional beings. So, it remains that properties, both absolute and relative, are not necessary condi-tions for the individuation of a fictional being. What does replace properties, then?

In point of fact, Thomasson says that the appearances of a character in different literary works in different possible worlds do not affect the identity of such a character, "provided that its point of origin in the author's creative acts exists in that world, and that the literary works concerning that character may be

traced back to it" (Thomasson 1999: 39). Indeed, Thomasson clearly thinks that, insofar as a fictum rigidly depends on its creative act – taken as a certain mental particular – such an act is a necessary condition for it (Thomasson 1999: 39, 109). Yet if we push the train of ideas contained in the above quotation up to its extreme consequences, it seems that Thomasson should maintain that the creative act also is a metaphysically sufficient condition, namely a condition whose satisfaction in any world enough is in order for something to be the fictum in question in that world. For suppose a world in which a certain author, Collodi for instance, performed the same creative act he had in our world when generating Pinocchio, by writing (if not merely thinking) 'Mastro Cherry happened to come across a thing'. Suppose moreover that Collodi, after having written (or simply thought) the above sentence, instead of proceeding in the story as he actually did, simply stops writing and just gives up his literary project. If we followed Thomasson's account, we should conclude that Pinocchio – our Pinocchio – also exists in that world. Yet this idea strikes me as utterly implausible. For argument's sake, let us acknowledge that Collodi's thought would be directed upon a certain intentional object, and even that such an object would be a fictional entity (admittedly, a rather tiny one): the thing to which is ascribed the property of being come across by Mastro Cherry. Yet it seems to me that such a tiny entity could hardly be the same as our Pinocchio.

Perhaps Thomasson would reply that Collodi's creative act in such a world is not the same mental token as the creative act actually generating Pinocchio. For what Collodi actually wrote (and thought) was not the previous sentence, but rather the following one: 'Mastro Cherry happened to come across a piece of wood that wept and laughed like a child'. In actual fact, Thomasson is silent as to how the particular creative act of thought must be individuated. Since she says that one and the same character may appear in different works in different possible worlds, provided that the creative act is the same, one is legitimized in supposing that the mental particular which constitutes such an act is individuated regardless of its content (so that it may have different contents in different worlds). Yet even if we conceded the above reply to Thomasson, and we got a fictional entity to which is ascribed the mere property of being a piece of wood come across by Mastro Cherry that weeps and laughs like a child, we would still be very far from having our

Pinocchio. Is there any reason as to why this tiny fictional entity, unlike the previous one (i.e. the thing merely come across by Mastro Cherry), should be the same as Pinocchio?

Let me just add that in raising these doubts I do not presuppose a realist conception of possible worlds à la Lewis. My issue is not that we first have to identify a possible world so that only then we can ask who is a certain individual in such a world. Rather, my point is to check whether certain properties are or not essential for a certain individual. If a certain property is essential for an individual, then, if in a certain world that individual does not instantiate that property, it does not exist in such a world (cf. Brody 1980: 84; Forbes 1986: 3). As a result, we can imagine many possible worlds in which it fails to have its accidental properties, but not worlds in which it fails to have its essential properties. So, we can peacefully imagine worlds in which a fictional object, say Pinocchio, is not loved by children. Yet to imagine a world in which Pinocchio fails to have the in-the-story-properties it actually has is to imagine a Pinocchio-less world.

Thomasson might rejoin that she just has the opposite intuition, namely that the tiny entity in question is Pinocchio, for it is conceived in the same creative act as in the actual world. Yet the problem here is not a mere clash of intuitions, as for instance in the case of biological individuals, where some – the Kripkeans – say that identity of origin suffices for individual identity, and others deny that it does (cf. Forbes 1980). The point is that if ficta have to be artifacts in any interesting sense of the word, they have to be constructed entities. As a result, identity of origin cannot suffice at all, for the constructive character of a fictional being cannot reside at all in such an origin (similarly Sutrop 2001: 137-138). Otherwise, any object which is brought into existence by its being thought of in a certain intentional act – as Brentanian intentional objects are (Brentano 1924: 88) – would eo ipso be an artifactual entity.

In point of fact, I think that Thomasson would share these doubts. Following her idea that a fictional entity not only both historically and rigidly depends on its author's creative act, but also both constantly and generically depends on some literary work or other in which it is spoken of, she says that what is sufficient for a fictum is not only the creative act, but also some or other literary work about it: "an author's creative acts and a literary work about the character are jointly sufficient for the

fictional character" (Thomasson 1999: 39). Yet this move actually fares no better. First of all, appealing to literary works in the individuation conditions of a fictional character raises the question of how such works have to be individuated, and we have to be sure that there is no circularity in such an individuation.[7] More problematic, however, is that such an appeal just pushes the original problem one step backwards. Suppose that Collodi, after having written (or thought) the above sentence ('Mastro Cherry happened to come across a piece of wood that wept and laughed like a child'), started writing a story that from that point onwards were utterly different from our Pinocchio. True, this time we would have a large character at our disposal. Yet again it seems to me that such a character would hardly be the same as Pinocchio. Of course, this time we would have not a merely thought-of, but a really constructed entity. Yet the construction involved would be utterly different from the actual one. In the case of concrete artifacts, if, identity of plans notwithstanding, the construction had given rise to a completely different thing from the one actually built, we will speak of different individuals. Why should matters be different in case of abstract artifacts?

Thomasson would probably reply that by "(jointly) sufficient conditions" for a fictum's individuation she does not mean metaphysically sufficient conditions, as I have done all along, but merely factually sufficient conditions, namely conditions whose satisfaction in a certain world (typically, ours) is enough in order for something to be the fictum in question in that world. Yet I strongly doubt that creative act + maintenance in some literary work or other provides even factually sufficient conditions. Suppose that after its completion, the original and unique textual copy in which a certain act of story-writing has been performed disintegrates; nor its author can talk about her literary project to anyone else, for she is disintegrated as well. In such a situation, in fact, there would be no way out of a certain make-believe practice of story-writing. Hence, it would be improper to say that that practice has been crystallized into a piece of fiction. *A fortiori*, in such a case, no ficta would be generated out of that practice.[8]

---

[7] For how such a problem affects Thomasson's account, cf. Iacona – Voltolini (2001).

[8] Of course, speaking of disintegration of a written copy as well as of its writer is a way of dramatizing things. To make my point, it would be enough if, after the

3. So far, we have seen that both the neo-Meinongian and the artifactualist conception of fictional entities face the same problem. Even if they managed to provide necessary conditions for fictional characters, they definitely fail to provide sufficient conditions. This prompts the question of whether the two perspectives are really mutually exclusive, as their followers seem to think. To my mind, there is room to answer this question negatively. Once both conceptions are suitably adjusted, one may say that both succeed in providing necessary as well as jointly sufficient conditions for fictional entities. Defending this idea amounts to maintaining what I would like to call the syncretistic theory of ficta.

According to such a theory, ficta are compound entities, made both of a set-theoretical and of a game-theoretical part. Each of these components is thus a necessary condition of a fictum; both are also jointly sufficient conditions.

The former part is easy to tell. As in the neo-Meinongian theories, there is a set involved in a fictional entity as one of its components, hence as a necessary condition of its individuation; namely, a set of properties. Unlike the neo-Meinongian theories, however, I do not say that these properties are those mobilized in the relevant literary work. For I want to leave the question of the identity of a literary work unsettled. I prefer to say that these properties are those mobilized in the relevant game of make-believe that underlies the constitution of the fictional entity in question. In the simplest cases, the make-believe game is the one a certain author engages himself while writing a certain text. In more complex cases, it may be a game played with oral words, involving more than one participant in its institution. (Mythological entities come out of games of this sort.)

The latter part, sufficiency, is perhaps a bit more complex to establish. First, artifactualist theories seem to hold a quite similar thesis when they say that a fictional entity depends on a creative act. Yet by saying that a creative act is (as well as a property set) a component of a fictum, I mean something stronger than the fact that the fictum depends on such an act, namely, that such an act is relevant for the individuation of that fictum. As Fine has nicely shown (Fine 1995 and elsewhere),

completion of the act of story-writing, the author had both abandoned the relevant copy and forgotten having written it.

mere dependence relations do not suffice for individuation. To be sure, a fictum rigidly depends on (the creative act) of its author; but it also rigidly depends on that author's mother. For not only it is true that necessarily, if that fictum exists then its author exists, but it is also true that if that fictum exists, then that author's mother exists as well. Yet the author's mother is clearly irrelevant for the individuation of that fictum.[9]

Moreover, unlike Thomasson I do not think that the creative act is a mental particular, but rather that the creative act is the relevant typological part of a certain make-believe game. For if we want the fictum to be a constructed entity, there must be a practice that supports its constitution, not a mere act of thought. I take the practice in question the relevant part of the whole make-believe game in which originally the fictum in question is spoken of, i.e. the game where the properties constituting the set-theoretical component of that fictum are mobilized.[10] Such a practice is moreover to be meant as a type, namely, something which may be instantiated by different (durative) event-tokens.

Both conditions are necessary; neither is by itself sufficient. The set-theoretical component is not sufficient, as both the 'no ficta' and the 'many ficta' problem show. But the game-theoretical component is also not sufficient. As I said above, if after its completion the original and unique copy of a certain work is disintegrated, and its author cannot talk about his literary project to anyone else, for he is disintegrated as well, then no ficta would be generated out of that fictional practice.

Nevertheless, these conditions are jointly sufficient. The make-believe practice-type mobilizes a certain set of properties. Yet, as we just saw, the practice itself is not enough. In order to have a certain fictum, that set must be added to that practice itself. (Typically, this is what happens when we step outside the practice and we start speaking about it: see later.)

This is why in the Menard case two ficta are involved. For although just one property set is mobilized, two distinct types of practices are at play. These practices are indeed different in type, for, although they mobilize the same properties, they are

---

[9] In actual fact, I think that Thomasson points in the same direction when she says that the creative act is a necessary condition of a fictum.

[10] In principle, Thomasson is not against this idea. For she admits (Thomasson 1999: 7) that the creation of a fictum might be, as she says, "diffuse".

causally disconnected.[11] In this sense, we may take a fictional entity at least as a 'many-one' set-correlate, rather than as a 'one-one' correlate as in some versions of neo-Meinongianism, for different fictional beings may correspond to one and the same property set.[12]

Two more questions have to be addressed. The first is: why does a fictum have to be taken as a correlate of a set rather than of a kind? As we have seen before, the neo-Meinongians hold either that a fictum is a set or that it is a kind. Why in this syncretistic approach do I opt for the former?

My answer involves saying something more about the practice (type) that supports the constitution of a fictional entity. In such a practice, what is mobilized is precisely a certain amount of properties: normal, ordinary properties. In actual fact, as Castañeda originally pointed out (Castañeda 1989a: 186-187), nobody doubts that the predicative terms which are involved in such a practice (whether oral or written) have the same meaning as they have outside such a practice: that is, they denote the same properties. Now, it has been traditionally said that unlike the games of make-believe involving real concrete individuals, the analogous games allegedly involving fictional individuals are existentially creative rather than conservative. Instead of taking a real concrete individual and make-believing of him that it has certain properties, in existentially creative games the actors of the game makes-believe that there is a certain concrete individual and that it has certain properties (Evans 1982: 358). Yet, under the assumption that a make-believe game mobilizes a certain amount of properties, the distinction between existentially conservative and existentially creative games of make-believe can definitely be attenuated. For, given the above as-

---

[11] A causal-intentional link between pretenses is definitely a necessary condition for their type-identity, as both Evans (1982: 362, 368) and Walton (1990: 403) suggest in general. By itself, it is however not a sufficient condition. Cf. Thomasson (1999:68-9).

[12] In this sense, my proposal remains different from Zalta's recent attempt at combining make-believe practices and neo-Meinongian abstractionism. According to Zalta, a fictum is an abstract Meinongian object that originates in a story and is thereby grounded in a certain make-believe practice. Cf. Zalta (2000: 127-128, 138-141). According to this position, although a fictum depends on such a practice, it is not a set-correlate, not only qua 'one-one', but also at least qua 'many-one' correlate. I say 'at least' for ficta may also be more naturally conceived as 'many-many' set-correlates: distinct practice-types may match not only the same, but of course also different, property sets.

sumption, making believe that a certain individual has certain properties is then the same as making believe, of a certain amount of properties, that a certain concrete individual has them. But then one is legitimized in taking an existentially creative game precisely as a game in which, of a certain set of properties, its actors make-believe that it is a concrete individual having those very properties.[13] So, a fictum must be composed by a set of properties rather than by a Platonic kind. For it is precisely on such a set that the relevant make-believe practice (type) operates.

The second question is: how can the relevant practice in a make-believe game be ontologically constitutive, or better, support the constitution of a fictional being? My answer is: look at the cases in which we resume an original practice, by playing a (possibly enlarged) game of make-believe involving such a continuation. This is what we typically do when we write a new body of literature that is intended as a continuation, or as a variation, of a previous narrative game. Literary cycles, as well as mythological stories, work precisely like this. Now, whenever we resume an original practice we change the fictum involved in the original practice, by adding properties that the original fictum did not have. Resuming a practice is not a neutral matter. For as soon as we resume it by uttering e.g. that a certain fictum is not only F,G,H... but also I,J,K, we mobilize a larger set of properties so that, at the end of such a resumption, we end up with an utterly different character.[14]

In looking at the same matter, Thomasson (1999: 67) says that in such cases it is just the same character that is involved all along. By linking one's self intentionally back to an original creative act, she claims, a new author (or possibly even the same author) simply ascribes new properties to the very same fictional being that has been generated in that act of creation. To be sure, the new part of the make-believe practice starts with pretending that the same concrete individual as the one

---

[13] To be sure, those taking part in the practice are not aware of making-believe of a set of properties that it is a concrete individual with those properties. Yet this is not a problem insofar as – pace Currie (1990) – making believe it is not a propositional attitude, but rather a use of props in an imaginative activity, as Walton (1990: 67) maintains; hence, something that does not automatically require first-person authority.

[14] In point of fact, we can even take the resuming practice as constitutive not of a 'larger', but of a merely different fictional entity, the one characterized simply by the properties that are mobilized in this practice.

that the previous part of the same practice made-believe it existed has different properties. Yet this does not mean that the new part of the make-believe practice manages to offer new characterizations of the very same fictional object. For, if as we have seen before a make-believe practice does not manage to support a mental act in order to get a fictional object, it cannot help such an act in order to get the same fictional object either. In actual fact, Thomasson's claim shows again her dispensing with ascribed properties in the individuation of a fictional being. As for her it is the very same x which in a certain possible world is described in a certain way (in a certain literary work) and in another possible world is described in another way (in another literary work), provided it is generated in the same creative act, for her it is again the very same x which in the very same world is described in a certain way (in a certain literary work) at time t and described in another way (in another literary work) at time t', provided both descriptions effectively trace back to the same creative act. [15] Yet as this theory sounded implausible in the modal case, it seems to me implausible in the temporal case. For, as I said before, adding maintenance of a fictum within some literary work or other to a creative act (conceived in terms of a certain particular act of thought) do not provide a jointly sufficient condition for the identity of a fictional being. Sic stantibus rebus, if we disregarded ascribed properties in the individuation of fictional beings, the upshot would be that we would have constituted no ficta until we were sure that the relevant make-believe game purportedly concerning such entities had come to an end. As no such assurance can be provided – for everybody can at any moment extend a make-believe game – we would end up with the ontologically disastrous result that we would have at our disposal not even a fictum up to the end of the world!

4. At this point, however, a further, though related, doubt arises. What proves that, although admittedly along with a set-theoretical element, the make-believe practice is ontologically constitutive? Traditionally, those that have appealed to games of make-believe in order to account for fictionality have main-

---

[15] I say 'effectively' for, as Thomasson herself acknowledges (1999:68), the intention of talking again of the same character in a new narration is merely a necessary, but not a sufficient, condition for intertextual identity of a character.

tained that whenever it is a question of fictional entities, make-believe is all that there is. That is, there is just the discursive practice of making believe that there are such entities, but there are no such entities in any genuine sense of the term 'there are' (cf. Evans 1982, Walton 1990).

To be faithful to believers in mere make-believe, their ontologically dismissive attitude does not absolutely mean that whenever we speak of fictional entities, we are engaged in a make-believe game. There is a distinction everyone taking part to the discussion wants to retain, namely the distinction between conniving and non-conniving uses of sentences allegedly involving fictional beings.[16] One would have a really ad hoc theory if one maintained that, unlike all the other cases where games are involved, as far as make-believe games are concerned there is no distinction between speaking in the game and speaking of the game, between speaking in fiction and speaking of fiction.

Put another way, a sentence allegedly involving fictional beings has not only mock-, but also serious, truth-conditions for believes in mere make-believe as well. The point is that believers in mere make-believe try to account for serious truth-conditions allegedly involving fictional beings in an ontologically non-committal way. Generically speaking, the idea is that the serious truth-conditions of one such sentence are the same as the conditions of its fictional truth. In other words, such a sentence is seriously true iff it is fictionally true. Using asterisk notation for signaling fictionality, we can say that 'p' (when non-connivingly used) is true iff *'p' is true* (for this general formulation cf. Crimmins 1998: 2-8). In a make-believe approach to fictional truths, moreover, to say that a sentence is fictionally true means that if one mockly asserts that sentence in a certain make-believe game, one makes-believe of herself in that game that one is speaking truly; more succinctly, a sentence is fictionally true just in case it is correctly attributed truth within the pretense (Walton 1990: 400; Crimmins 1998: 4; Evans 1982: 365).

---

[16] For this terminological pair, cf. Evans (1982: 365-366). Currie (1988: 471, 477; 1990: 146-162) draws a similar distinction between fictive and meta-fictive uses of fictional sentences. Yet this distinction is not the same as the previous one, for in Currie's account both the fictive and the metafictive sentential use have – admittedly distinct – serious truth-conditions.

Now, there is an already wide debate in the literature on whether the pretense approach gives adequate truth-conditions for any case of non-connivingly used sentences. As many have maintained (cf. Stanley 2001), the answer to this question is negative. I believe that this answer is correct. For any particular non-committal paraphrase of such sentences in that use that can be provided within such an approach, I think that it is possible to find a convincing counterexample. Yet I do not want to take on this debate, for I am here mainly concerned with ontological rather than semantic issues. So, I limit myself to raising a general problem to the very truth-conditional strategy this non-committal approach intends to pursue.

Let us go back to the alleged equivalence according to which when non-connivingly used a sentence is seriously true iff it is fictionally true. I have no qualms as to the right-to-left direction of the equivalence: being fictionally true is definitely a necessary condition for a sentence to be seriously true. All the game situations whose players engage in the same pretense as the original player of the game (possibly, a writer) make those players utter fictional truths.[17] Moreover, the original player of the game utters (or anyway generates)[18] those very same fictional truths by stipulation. For in mock-asserting (directly or indirectly)[19] anything within that game, she eo ipso make-believedly speaks truly.[20] Now, it may be the case that there are no further participants to a certain make-believe game pretending the same as its original player. Yet there must at least be such an original player in order for the relevant fictional truths to obtain. Now, if there were no such truths, there could hardly be the corresponding serious truths either. How could it seriously be the case that Pinocchio is a piece of wood if nobody, ultimately not even Col-

---

[17] If I understand Walton correctly, for participants to further instances of a make-believe game to engage in the same pretense as the players of its original instance means that those participants comply with the prescriptions to imagine that are given in such an instance. As Walton (1990: 51) puts it, what I call further instances of a make-believe game are different make-believe games that are authorized by the original make-believe game, insofar as both rely on the same props.

[18] I put this clause in order to take into account that there are not only uttered, but also implied, fictional truths.

[19] See previous footnote.

[20] In this sense, as far as the original players of a make-believe game are concerned there is no difference between making-believe and making up a story, as Deutsch (2000) instead maintains. For one makes up a story insofar as what one makes-believe is eo ipso a fictional truth.

lodi, had make-believedly spoken truly by mock-asserting (directly or indirectly) the corresponding sentence?

Yet there are problems with the left-to-right direction. For being fictionally true hardly is a sufficient condition for a sentence to be seriously true. Saying that a sentence is fictionally true does not eo ipso mean that the sentence is also seriously true. Assigning a sentence a fictional truth-value has no import at all on whether that sentence also has a serious truth-value. To assign it a serious truth-value we have to step outside pretense, hence outside also the very practice itself of mock-speaking truly. If we appealed to the (admittedly dangerous) terminology of fictional worlds, we might put things in this way: the fact that a sentence is true in a fictional world has no import at all on whether that very sentence is also true at the actual world.[21]

We can clearly see what the problem is if we consider the admittedly analogous case of dreams.[22] In a dream, its subject utters certain sentences. At least some of these sentences are oneirically true: in the dream, things stand as these sentences present them to be. Yet the fact that these sentences are oneirically true does not eo ipso mean that they are seriously true. Real truths are not acquired for the simple fact that one dreams sentences as true.

In my ontological framework, I can easily account for this problem. In order for a non-connivingly used sentence to be seriously true, it does not suffice that it is fictionally true. For one and the same sentence moves from being fictionally true when connivingly used onto being seriously true when non-connivingly used insofar as its singular terms move from pretending to refer to (concrete) individuals onto really referring to other individuals, namely the ontological compounds in which ficta consist.

As I said at the very beginning, this shows that my theory fulfils an even more conciliatory aim than that of combining neo-Meinongianism with artifactualism. For if I am right, then the

---

[21] Note that this problem would not affect those that were tempted to analyze the non-conniving use of a sentence 'p' à la Lewis (1978), i.e. in terms of an intensionalist paraphrase of the kind 'in the story S, p'. For within the intensionalist approach, a sentence of the latter kind, hence the paraphrased sentence in its non-conniving use, is true at the actual world iff the embedded sentence 'p' is true not in, but at the world of the fiction. However, I cannot deal with the intensionalist approach here.

[22] Walton (1990: 43-50) has precisely maintained that dream sentences have to be treated in the same way as fictional sentences.

make-believe approach to fiction is compatible with a committal approach such as the one presented above. Indeed, insofar as ficta figure within the serious truth-conditions the connivingly used sentences possess when they move onto being non-connivingly used, my committal approach only regards non-connivingly used sentences, whereas the make-believe non-committal approach limits itself to regarding connivingly used sentences.[23]

BIBLIOGRAPHY

Brentano, Franz. 1924. *Psychologie vom empirischen Standpunkt* (trans. A.C. Rancurello et al.). Band I. Hamburg: F. Meiner (London: Routledge 1995[2]).

Brody, B.A. 1980. *Identity and Essence*. Princeton: Princeton University Press.

Castañeda, Héctor-Neri. 1989a. 'Fiction and Reality: Ontological Questions about Literary Experience' in *Thinking, Language, and Experience*. Minneapolis: University of Minnesota Press: 176-205.

Castañeda, Héctor-Neri. 1989b. 'Thinking and the Structure of the World' in *Thinking, Language, and Experience*. Minneapolis: University of Minnesota Press: 235-261.

Crimmins, Mark. 1998. 'Hesperus and Phosphorus: Sense, Pretense, and Reference' in *The Philosophical Review* 107(1): 1-47.

Currie, Gregory. 1988. 'Fictional Names' in *Australasian Journal of Philosophy* 66(4): 471-488.

Currie, Gregory. 1990. *The Nature of Fiction*. Cambridge: Cambridge University Press.

Deutsch, H. 2000. 'Making Up Stories' in Everett, Anthony and Thomas Hofweber (ed.) *Empty Names, Fiction and the Puzzles of Non-existence*. Stanford: CSLI Publications: 149-181.

Evans, Garreth. 1982. *The Varieties of Reference*. Oxford: Clarendon Press.

Fine, Kit. 1995. 'Ontological Dependence' in *Proceedings of the Aristotelian Society* 95: 269-289.

Forbes, Graeme. 1980. 'Origin and Identity' in *Philosophical Studies* 37(4): 353-362.

Forbes, Graeme. 1986. 'In Defense of Absolute Essentialism' in *Midwest Studies in Philosophy* 11(1): 3-31.

Iacona, Andrea and Alberto Voltolini. 2001. Review of Thomasson (1999) in *Dialectica* 55: 379-382.

Kripke, Saul. 1973. *The John Locke Lectures 1973*. Unpublished typescript.

Lewis, David. 1978. 'Truth in Fiction' in *American Philosophical Quarterly* 15: 37- 46.

Meinong, Alexius. 1916. *Über Möglichkeit und Wahrscheinlichkeit*. Leipzig: Barth. Rpt. in Gesamtausgabe Bd. VI, Graz: Akademische Druck- und Verlagsanstalt 1972.

Meinong, Alexius. 1960. 'The Theory of Objects' (trans. Levi, Isaac, D.B. Terrell and R.M. Chisholm) in Roderick Chisholm (ed.) *Realism and the Background of Phenomenology*. New York: Free Press: 76-117. (Originally 'Über

---

[23] Zalta (2000: 141-144) provides another attempt to combine make-believe and genuine reference to ficta. Yet his ontological perspective is different from the one presented here. Cf. fn. 12.

Gegenstandtheorie' in *Untersuchungen zur Gegenstandtheorie und Psychologie* (ed. Alexius Meinong). Leipzig 1904. Rpt. in Gesammelte Abhandlungen, Gesamtausgabe Bd. II, Graz: Akademische Druck- und Verlagsanstalt 1971: 481-535.)

Parsons, Terence. 1980. *Nonexistent Objects*. New Haven and London: Yale University Press.

Salmon, Nathan U. 1982. *Reference and Essence*. Oxford: Blackwell.

Salmon, Nathan U. 1998. 'Nonexistence' in *Noûs* 32(3): 277-319.

Sutrop, Margit. 2001. Review of Thomasson 1999 in *European Journal of Philosophy* 9: 135-139.

Stanley, J. 2001. 'Hermeneutic Fictionalism' in Wettstein, Howard and Peter French (ed.) *Figurative Language* (Midwest Studies in Philosophy 25). Oxford: Blackwell: 36-71.

Thomasson, A.L. 1999. *Fiction and Metaphysics*. Cambridge: Cambridge University Press.

Van Inwagen, Peter. 1979. 'Creatures of Fiction' in *American Philosophical Quarterly* 14(4): 299-308.

Voltolini, Alberto. 1994. 'Ficta versus Possibilia' in *Grazer Philosophische Studien* 48: 75-104.

Walton, Kendall. 1990. *Mimesis as Make-Believe*. Harvard University Press, Cambridge (Mass.).

Zalta, Edward N. 1983. *Abstract Objects*. Dordrecht: Reidel.

Zalta, Edward N. 2000. 'Pretense Theory and Abstract Object Theory' in Anthony Everett and Thomas Hofweber (ed.) *Empty Names, Fiction and the Puzzles of Non-existence*. Stanford: CSLI Publications: 117-147.

KAREL THEIN
Menard, the *Quixote*, the Indiscernibles

Abstract: The figures of exhaustion, whether real or pretended, enable Borges to merge logic and human condition. The resulting blend situates human existence at the intersection of metaphysical and cosmic themes. Exemplifying Borges' use of the whole philosophical systems as literary tropes, *Pierre Menard, Author of the Quixote* belong to those stories that present us with an *object* (a novel, a coin, an artificial language) whose ontological status and the latter's implications can be studied more or less separately from their literary context. These objects invite the elaboration of a possible ontology: more specifically, Menard's effort at writing the *Quixote* revives the ancient and early modern discussions about the indiscernibles (non-distinguishable entities) and, ultimately, sketches a new kind of metaphysical history where the contingent and the necessary would merge into one.
Keywords: ontology of literary work; literature and philosophy; metaphysics of the indiscernibles

Borges' combinatorial style relies on a strong impact of a rather small number of poor and openly repetitive figures where oxymoron seems to enjoy a particular privilege. This style, whereby metaphysical problems change into a confused bundle of frivolous and superficial paradoxes, claims a melancholy and haunted knowledge which follows from the recognition that repetition and combination of a deliberately restricted inventory of elements may not be proper of the literature, but of the real universe, a work of an unknown and long-time forgotten author. Borges' lullabies of a tired cosmos thus remind their human readers of their own place in the impenetrable yet *probably* large scheme of things.

This allegedly simple knowledge helps to purify literature of the residual sense of humor (despite the undeniably funny effect this has on the reader) and to put limits to every attempt at analysis. In its usual and theoretical sense, the latter seems rather superfluous since it is structurally identical to (or at least not distinguishable from) its very object. Still it can reformulate some rules of the game, of the cosmic as well as literary lottery,

without of course being able to determine the meanings that will result from its next round. Borges seems to imply that the old universe, composed as it is of the endless series of past and future variations, contains both literature and literary science as two forms of one and the same weariness. Time, after all, erases their traces regularly yet involuntarily, just like the cow's tail that drives away the annoying and buzzing flies.

The figures of exhaustion, real or pretended, enable Borges to merge logic and human condition.[1] Lives of men are not tragic and, as times go by, they lose their comic side too. The life itself in its doubtful individual manifestations appears as a series within a series, in other words as more variations on the main cosmic themes. In most lifetimes, Borges seems to imply, the inventory of possible occupations runs out quickly and a low, boring form of repetition steps in. However, some carefully selected and apparently (or perhaps truly) non-sensical tasks may fulfill a man's life with voluntarily constructed repetitions as a means of avoiding the involuntary ones. The list of Menard's 'visible work' is a perfect illustration of this scheme: except for two literary works proper (both of them being sonnets), the focus is invariably on combinations and transpositions. In this respect, Menard clearly belongs among those Borgesian heroes who live an episodic existence consisting in a painstaking production of intellectually intriguing games. In many other cases Borges' short stories hint at a possibility that these games, ultimately trivial from the human point of view, hide some clues to cosmic redemption or survival. This, however, is not the case with Menard. The latter is the opposite of the second type of a Borgesian hero, an uneducated, savage genius, a gaucho whose untimely death prevents forever any intellectual awakening and achievement (the memory-genius Fuñes comes immediately to one's mind, but there are many others and perhaps better examples).

Now it is perhaps not without interest to remark that, by contrast to Menard, it is Cervantes who would fit quite well into this second category, including his capacity to produce (truly 'single-handedly') a great narrative that became one of *the* Modern novels. For a reader of Borges, Menard's choice of *Quixote* is most interesting precisely by its willingness to confront,

---

[1] I would like to believe that this claim differs from the once famous position of John Barth on the 'literature of exhaustion' as expressed in his eponymous essay. See Barth 1977.

by means of a necessarily detailed linguistic construction where not a single iota can be changed, the originally untamed source that had produced, on its own, such an ironic and monumental *chef d'œuvre*. No doubt it is a measure of Borges' success that his very short story manages to retain *both* the irony *and* the monumentality, not without transposing them on a different scale, yet preserving much of their original flavor. It is not only that the literary *amateur* and *bricoleur* Menard is indeed a Quixote in our general sense of the word; rather, it is the short story's own ability to appropriate the monumentality by suggesting (rather than admitting) its hero's necessary failure.

Were we to start a formal analysis of this issue, it would, I suspect, boil down to the basic tension between repetition and narration, and to various ways narration can be constructed out of repetitions. Still it is more important to emphasize that Borges himself (again by contrast to Menard) seems to have little patience with formal schemes proper: he loves to enumerate them, but not to carry them out. One should also not forget that Borges' style, with its quasi-romantic chant of indifference and paradoxical fragments of monumentality, opposes quite vigorously the modern literature's 'existential' temptation of subjectivity and introspection. To Borges, any inwardness (which may be just boring in any case) seems accessible without its falsification. And to a false psychological depth, a real cosmic or even theological surface is to be preferred. Shining out of the carefully polished surfaces as that many sketches of universes created by dreaming gods, Borges' most memorable and hence often-quoted sentences are summaries. Having eliminated the last traces of week psychology, Borges can return, time and again, to his favorite antithesis of human reason and redemption: 'What will be left for you is days and nights, the stereotypes of thought, everyday life and the world' (*Blue Tigers*).

This wonderful removal of wonder (an element of a talkative falsity) would seem to be a gold mine for philosophical interpretation. On a closer look, however, the relative paucity of real and close readings of Borges by philosophers prevails. The very medium of a short story has probably a lot to do with this situation, and also the correlative fact that Borges himself keeps preempting the philosophical rereading by abbreviating the whole philosophical systems and using them as that many skeptical

tropes.[2] It is then not surprising at all that we find *Pierre Menard, Author of the Quixote* among the stories that attracted the most successful philosophical efforts: together with *Zahir* or *Tlön, Uqbar, Orbis Tertius*, it presents us with an *object* (a novel, a coin, an artificial language) whose ontological status and the latter's implications can be studied more or less separately from their literary context. These objects invite the elaboration of a *possible* ontology (even the bizarre *zahir* does) – by contrast to stories like *Blue Tigers* that place at their center an *impossible* object (note that Borges does not hesitate to mention extravagant yet possible object and the impossible ones within one and the same story, like in *Zahir* where a coin with only one side is mentioned).

While the ontologies of Menard's *Quixote* and of the latter relation to the *Quixote* of Cervantes represent legitimate and important philosophical concerns, it is still true that they are somehow obliged to escape the framework of Borges' dealing with philosophy. More exactly, they need (and have the right) to side-step his ironic double strategy, which consists in the (metaphysically speaking) *realist* treatment of many nouns, especially names on the one hand, and the *nominalist* epistemology and amused skepticism on the other hand. This strategy seems often to enable Borges to appropriate philosophy in order to make it simply part of literature. To this end, he approaches the philosophical themes either from the non-human perspective of eternity, or through the threat of an intimate irrationality of a world whose incommensurable dimensions are as if endless and created at the same time (hence the proximity to Lovecraft: see, most remarkably, the description of 'the hideous City of the immortals' in the short story *The Immortal One*). The skeptical effect of this strategy is just that: an effect rather that a deliberate inquisitive stance. Which is also why it connects so well to the more or less arbitrary pairing of various philosophical theses with all sorts of objects – what we got is a sort of *historia naturalis* of human spirit, not its 'true' genealogy.

Enumeration thus replaces any real philosophical method. It is Borges' way of being simultaneously repetitive and laconic, of talking about so much while saying, after all, so little. This sty-

---

[2] Here I will not deal with Borges' relation to skepticism and to his use of some skeptical strategies in order to attain a sort of literary *ataraxia*. My use of the noun 'skepticism' and the adjective 'skeptic' is thus a very loose one.

listic feature of his short stories, including *Menard*, is what belongs fully to the realm of literary, not philosophy. Moreover, within this realm, it helps to put aside the traditional literary ally of philosophy, the novel. Short story is what Borges opposes to the novel quite deliberately, as a crystallization of the skeptical effect whereby even the much-celebrated laughter of the novel is put in its place.[3]

In much the same vein one could remark that even the most successful readings of *Menard* can still fuel the doubts whether, at the end of the day, there is or can ever be such a thing as a true and non-derivative ontology of the artworks as works of *art*. Not only can one suspect that the intriguing and as if inexhaustible nature of the relevant artworks may reflect a lack of any determinate ontology in the first place. It is equally possible to evoke again the already mentioned fact that, in the case of most of Borges' stories, interpreters apply ontological categories in order to relate the latter to either the broader issue of fictional objects or to the narrower issue of particular artefactual nature of the objects under (literary) description. I have no doubt about the validity of results thus obtained; still these results often confirm that Borges' counterintuitive yet not conceptual objects are most intriguing precisely as *objects* and not, more specifically, as *contents* of the works of art.[4]

Be that as it may, Borges' procedures are asymmetrically opposed to any effort at explaining literature from this or that philosophical position. In their own way, they testify to the fact that no matter how intellectually minded its author may be, a literary text (regardless its genre and subject matter) has neither the nature nor the structure of a problem.[5] At the same time, they confirm that the closer a literary text comes to *appear* to possess such a nature and structure (by evoking classical phi-

---

[3] Borges' relation to the novel is of course much more complex and cannot be analyzed here. Suffice it to say that, even within the literary realm, his work happens to be seen today as generative of some new novelistic adventures – like, for instance, those of Roberto Bolaño.

[4] Since several contributions in this volume deal in some detail with this and related issues, I will return to them elsewhere.

[5] I cannot enter here the issue of how philosophical *problems* relate to philosophical *questions*. In any case, many people simply say that, by contrast to mysteries, both imply a possibility of an answer that would rely on roughly the same level of knowledge and intelligence. For the distinction between problems and mysteries and its inspiration in the work of Noam Chomsky see McGinn 1993; cf. Unger 2002.

losophical topics or having some apparently theoretical features), the bigger the challenger for its interpreters. This challenge is not one of identifying the sources and possible references but, more plainly, one of establishing its proper space alongside the eloquent work that may not *need* to be interpreted in the first place. There is little doubt that *Pierre Menard, Author of the Quixote* is a paradigmatic case of such a difficulty. To avoid some obvious traps which follow from this situation, I divide the following brief exercise in two modest parts: the first one is an attempt at a descriptive paraphrasis; the second one contains a fragment of a historical catalogue.

1. SHORT STORY, DESTRUCTION OF THE NOVEL

In *Pierre Menard, Author of the Quixote*, Borges' narrator is wavering and perhaps not very intelligent (though at moments brilliant); on the other hand, he is definitely full of admiration towards Menard. While mocking the whole circle to which the narrator belongs, Borges is not simply parodying his noble and conventional notion of being faithful to a dead acquaintance. Instead, he carries out a triple consequence of this notion: the defense of Menard's work including the latter's decision to *write* (rather than write *again*) the *Quixote*; the correlative emphasis on a real creation instead of imitation, copying or some other re-enactment; and finally the evaluation of Menard's dazzling and purely intellectual *Quixote* as neatly superior to the *Quixote* written, contingently, by some half-cultivated Spanish soldier.

The lack of trust in the allegedly pedestrian and savage 17th-century genius determines the first part of the text with its admiration for all sorts of intellectual and cultural refinement. This admiration finds its expression in the praise lavished on Menard of which the catalogue of the latter's work is an integral part. Consisting in nineteen different entries, the catalogue irresistibly suggests an image of a second-rate symbolist whose ostentatious intellectual asceticism makes for the lack of a real natural talent. Menard's true element is the check problems, antiquarian logical paradoxes, grammar and formally refined versifying. The catalogue itself lacks in a careful redaction; it has a very relaxed internal structure, which may be meant to confirm that Menard belongs to the actually not so small company of men with a real genius but no true talent (as we will come to understand, it is Cervantes' *work* that is taken for simply contingent; in Menard, the contingency afflicts the author).

In its basic general features, this situation is well known from many of Borges' texts: pointlessness implies the infinite, which then opens the door to a sort of pulp metaphysics. Metaphysics, by the way, becomes explicit once the narrator turns to Menard's opus magnum and its ambition at creating rather than re-creating, by means of logic and resurrection combined, the *Quixote*. Already at this point, before saying more about Menard's searching for the means of achieving his goal, one crucial implication of the whole text should be emphasized: in fact, the metaphysically amazing object of the story is not *Quixote* or the second-first *Quixote*, but its author, Pierre Menard. It might be ironic of Borges to let 'his' narrator to insist so strongly on Menard's originality (no simply copying or reproducing is involved; no actualization or anachronism either) while making only the moral (and morally distasteful) comments on the fundamental difference between the author and his work, and hence between the two authors.

This difference has nothing to do with the issues of personhood and authorship; it is simply a causal one. By 'author', I mean here the cause; by 'work', the latter's effect. Now surely Cervantes was the cause of the *Quixote*. If Menard becomes the cause of the *Quixote* too, he will only demonstrate that two different causes can have the same effect. Such an implication, no matter how much the second creation of an object as complex as the *Quixote* would amaze us, is a rather trivial one. In other words, since the *Quixote* is described as a text or a meaning, not as a physical object (a manuscript), it implies no metaphysical impossibility. By contrast, what could have been truly strange yet was explicitly abandoned in the course of the story is the identity or complete merging of the two causes. Menard, however, will *not* become Cervantes. Having considered and then rejected the option of recreating the conditions that *made* an otherwise unknown *someone* into Cervantes, Menard rightly assumes that he cannot try to *become* Cervantes since, metaphysically speaking, he would already had to *be* him from the very start of his enterprise. Yet were he Cervantes, his *Quixote* would be again *the other's Quixote*. In all, the metaphysics of Menard's enterprise self-destroys without the narrator taking much notice.

Menard's main problem is thus quite simple: the possibility of choice, which is the condition of achieving his full merging with Cervantes, is what automatically undermines such merging (the resulting identity would not be his own in any case). Menard's

acknowledgement of this problem shows through several passages of the text, including the narrator's almost theatrical praise of the heroic nature of Menard's failure. Here the heroism takes on a new meaning which is very different from, and indeed contrary to the classical epic tradition: Menard does not see his imminent death approaching, yet he prophesizes the grotesque fiasco of his soon to be forgotten work (in much the similar way, he is more interested in Achilles' pseudo-logical race against the turtle than in Homer's semi-divine Achilles). Menard's consciousness of his own failure is explicitly repeated in the text – and it would be easy to mistake his 'heroism' for a childish stubbornness. Hence the task of interpreting Borges' story without much supplementary theorizing yet so as to give it some sense independent of Menard's impossible success.

The simplest answer to such a task is to stick to the difference between Menard's point of view and the literary apparatus of Borges' short story. In fact, the impossibility of success in Menard's narrated case is the precondition of the narrator's *evaluative* attitude. Time and again, only the non-identity of both *Quixotes* allows for their different appreciation. The praise of Menard thus follows quite directly from his fiasco – which, importantly again, does not consist in the inability to write the *Quixote* (the narrator quotes the passage about history and truth in both literally identical 'versions'), but to *finish* or *complete* this voluminous novel. The fragmentary nature of the result is another enigma of the story: beyond its empirical dimension, is it an implicit judgment passed on the history and thus a revelation of what the quoted fragment expresses through other means? Or is it just a contingent fact that becomes a necessary part of the short story that, by the same token, implies the historical *destruction of the novel* that many readers take for the primary element of the Borgesian *œuvre*?

In both cases, *Pierre Menard, Author of the Quixote* hints at a rhapsodic revision of history whereby persons and events become just facets of some both larger and more detailed or refined mirror-play. The time's flow is experimentally divided into separate parcels that any author cannot help combining – voluntarily or not. By its very structure where much is left unsaid and what is said receives an illusion of understated importance, the short story can feed upon this assumption in a way that the novel cannot since it still is, on Borges' account, too deeply immerged into the river of time. Moreover, but this is well be-

yond the scope of any reading of *Menard*, the novel is a literary form originally tied to erotic adventures (in Antiquity) and to the coming of age (in Modern times). Borges, who holds both of these issues in delighted abhorrence, takes the short story for an obvious means of purifying literature. Nevertheless, even he seems to admit that the library and the bed are the opposite forms of infinite duplication, two distinct and forever distant poles of pleasure.[6]

Fed by the ruins of the novel, our short story closes with a paragraph that I cannot, despite a repeated effort, quite comprehend. The narrator claims that

> Menard (perhaps without wanting to) has enriched, by means of a new technique, the halting and rudimentary art of reading: this new technique is that of the deliberate anachronism and the erroneous attribution.

The assumption of an 'involuntary' invention is unavoidable here, since such a claim smells of the ultimate divorce between the art of writing and the art of reading. Yet this goes certainly against the grain of Menard's intention: not only did he not want to be a follower of Cervantes and even less his imitator, but he could hardly have accepted that Cervantes' *Quixote* could serve as a tool for our understanding of what *he* wrote. As he 'multiplied draft after draft' of his ('already extant') book, he intended to eliminate the space for interpretation rather than open up a new one.

2. A REMARK ON THE INDISCERNIBLES

> There are not, in the whole vast Library, two identical books.
> (Borges 1962: 82-83)

Seen through the lens of philosophy – thus from the outside –, *Pierre Menard, Author of the Quixote* is basically a variation on the old problem of indiscernibles that Borges puts to work on more than one occasion. Here as elsewhere, he reaches into the

---

[6] On this general conclusion, we can go back to the beginning and ask: why *Quixote*? Perhaps Borges is addressing this very question to himself – and perhaps that his autobiographical fictions contain for once a shadow of an answer: he pretends that the *Quixote*, read first in English (and only later in its 'poorer' Spanish version), made him realize the difference between the mother tongue and the language of literature. Can there be *here* a hidden core of Borges' understanding of history as impersonal autobiography of sorts?

repertory of philosophy both ancient and modern in order to borrow a theme that appears iterable in many different contexts. In our case, thus relatively to the issue of indiscernibles, the question is to know how to take Borges' description of Menard's clear failure: does the latter have some apriori explanation, or is it ultimately empirical in its nature? At the first sight, this second option seems too outlandish to be taken seriously, but we must not forget that Borges' perspective is not one of a philosopher or a logician. Writing against the background assumption of a fundamentally flawed and imperfect universe (thus on the borderline of a Gnostic position that philosophy simply cannot accommodate), our author indicates more than once that such a universe allows for illogical monstrosities without succumbing to them. From this perspective, in the unknown literary past, there *might have* been cases parallel to the one of Cervantes and Menard; the principle of indiscernibles *might have* been variously bent.

Together with *The Babylon Lottery* or *The Theologians*, *Pierre Menard, Author of the Quixote* ranges among the texts where this principle applies in a crystal-clear way. Its identification and a neat analysis of its application are due to Antoine Compagnon's book on the history and the structure of quotation (Compagnon 1979: 370-380). Compagnon relates Borges' stylistic procedures to the four basic principles of the philosophy of Leibniz: the law of non-contradiction, the principle of sufficient reason, the continuum and the indiscernibles or the principle that should exclude the existence of the absolutely identical individuals within the same universe. The interplay of these four principles, where the third and the fourth ones specify the second one, enables Leibniz to explain the world as a composition of various degrees of possibility and reality (in Descartes, says Compagnon, what is real follows from the choice among all the possibilities; in Spinoza, all the possibilities are realized).

It is certainly true that the world of Leibniz looks much more like a novel than a volume of short stories (after all, he was a great fan of Honoré d'Urfé's *Astrée*).[7] Yet its principles are indeed such that Borges can easily use them within the restricted space of a short story – and repeat this use with many variations in other stories. Thus he is allowed to do what Leibniz

---

[7] Among recent literary efforts, Neal Stephenson's monumental *Baroque Cycle* (2003-2004) takes up the challenge implied in this connection.

would hardly approve of: with a little help of not quite clear theology he can apply the principle of indiscernibles not just in different ways, but with different *results*. Whereas *Menard* seems to confirm the principle (in the realm of meanings at least, Menard's would be the richer and a more nuanced book), the short story *Theologians* is written in order to doubt its universal validity. In one word, it sets to imply that two human bodies *plus* minds *plus* the series of the predicates applicable to the former and the latter can be taken for something more than just similar: in principle, they may be taken for two perfectly identical beings (it must be said that the conclusion of the story is very vague and one is allowed to think about the two beings in question as a single one – instead of two non-distinguishable entities, they would be two overlapping parts of one whole).[8]

By contrast to contemporary analyses that develop Leibniz's use of the principle towards new and increasingly refined versions of possible worlds, Borges' references to Leibniz, though connected to the issue of *Quixote*'s possible meanings, belong rather to the ancient context where the whole debate originated as fundamentally a cosmological one. Although this feature is more apparent in *The Theologians*, it can still be profitably evoked in relation to *Menard*. Like elsewhere in Borges, issues of meaning are (or should be) indicative of the order or disorder of the physical universe.[9]

Although (as we are going to see) the most relevant ancient passage may come from Plato, the most direct link between Leibniz and the Ancients leads to the Hellenistic and Greco-Roman periods, more exactly to a polemical discussion between the Stoics and the Academical Skeptics. The shortest summary of the Stoic position is offered by Cicero in *Academica Posteriora* XVI, 85, during a longer discussion that summarizes differ-

---

[8] For more detail and the relation (or the opposition) to Leibniz's *Theodicy* see Compagnon 1979. It is not possible to list here the works relevant to the problem. For a concise discussion of possible worlds and possible individuals in the early modern philosophy see at least Wilson 1999.

[9] By contrast, I leave entirely aside the relevant discussions in the context of contemporary continental philosophy. One can think of Deleuze's concept of repetition or, on a different plane, of Foucault's reading of the *Quixote* that seems to offer an indirect critique of Menard's enterprise (and, by the same token, an unintended praise of its originality). Others have emphasized the possible links between *Menard* and the discussion of repetition in contemporary art theory (with respect to the later, Krauss 1981 is still an obvious point of reference).

ent appreciations of the puzzle of the cataleptic assent to an impression occasioned by sense perception: it is the problem of whether we can distinguish between a *visum verum* and a *visum falsum*. Leaving aside the complicated technicalities, the issue is one of identical impressions that produce mental confusion, and ultimately madness. Hence the question of whether there can be a physical correlate of such confusion: can there obtain two identical and yet individually existing entities? To simplify again, the Stoics say no: some entities can be very much alike but once we consider them with all their properties and predicates they are never truly identical. By contrast, a skeptic leaves the matter undecided: we have no neat external criterion that would help us to decide and experience itself is not sufficient to arrive at any firm conclusion; we can only rely on what is plausible (*probabile*) or persuasive (*pithanon*).

For a skeptic, it is thus impossible to agree with the Stoics that 'no hair or grain of sand is in all respects the same as another hair or grain.' This one-sentence summary does hardly justice to the combination of logical criteria with material ontology that lies at the heart of the Stoic argument, but it is more remarkable how the skeptic replies. Refusing to acknowledge this quoted contention as a simple truth, he appeals to the example of the arts: if we must always assume some distinction between natural being including humans, could a great artist not produce, say, a given number of indistinguishable statues? Perhaps he could; and yet it is still *possible* that an equally great master of the same art would be able to tell them apart (XXVI, 85-86).

If, then, we put the Stoics side by side with Leibniz, it is the skeptical counter-argument that seems to anticipate Menard's intention – despite the fact that they deny the possibility of its unequivocal fulfillment.[10] This anticipation follows precisely from the appeal to the arts, more exactly from a (purely hypothetical) suggestion that an artist could succeed in producing a naturally impossible object. In this oblique and unexpected way, the skep-

---

[10] The Stoics who, by contrast to Leibniz, reason on the basis of a temporal rather than logical plurality of worlds, are relevant here only insofar as their position implies a strict denial of any success whatsoever: defenders of a strong determinism across the successive universes, they would insist that the only and eternally repeated *Quixote* is the one written by Cervantes. – For the Stoic position on the eternal recurrence as described in Nemesius, *De natura hominis*, 309,5-311,1 (*SVF* II, 265), see Barnes 1978, who takes the doctrine for incoherent. An opposite view is defended in Long 1985.

tical objection brings in the issue of *deliberate* creation. However, in following this turn of the argument, we would be forced to quote the considerations produced well before the time of the skeptics by Plato.

Both *Republic* 10 and the *Cratylus* discuss the uniqueness of a deliberately created object. Despite the fact that the contexts are very different and that the former text speaks about the creation of a pattern which is to be further reproduced, whereas the latter deals with a reproduction of an already existing entity, in both cases Plato puts forward a hypothesis of a divine artist or craftsman who probably *could*, but *would not* create an object, be it an intelligible or a sensible one, which it would be impossible to distinguish from another object of the same kind. Plato, not unlike Borges, addresses the indiscernibility as a part of the wider issue of omnipotence and its logico-metaphysical *and* physical limits. It would be preposterous and useless to try to summarize both Platonic passages; suffice it to say that the one from the *Cratylus*, 432b4-c6, links the issue of divine power to produce an individual indistinguishable from another individual to the logically necessary *imperfection* of any given reproduction: once perfected, such a reproduction automatically looses its status of image and acquires an independent existence of its own. What Plato does not say in this text is whether the possibly divergent series of predicates of the two resulting objects possess some truly metaphysical relevance.[11]

I do not mean to pretend that Plato might have inspired Borges and even less that there is some sort of doctrinal resemblance between their respective texts. It is however important to recognize the possible mediation of Platonic themes through the distorted mirror of Gnostic turn towards the less potent (and less good) artistic divinities. This may furnish us with a link between the morally subverted Platonism and the original notion of literature as a field of sublunary adventures. In this connection, we must also not forget, as some interpreters surprisingly do,

---

[11] It may be useful to quote the passage in question (translated by C. D. C. Reeve): '*Socrates*: Would there be two things – Cratylus and an image of Cratylus – in the following circumstances? Suppose some God didn't just represent your colour and shape the way painters do, but made all the inner parts like yours, with the same warmth and softness, and put motion, soul and wisdom like yours into them – in a word, suppose he made a duplicate of everything you have and put it beside you. Would there then be two Cratyluses or Cratylus and an image of Cratylus? – *Cratylus*: It seems to me, Socrates, that there would be two Cratyluses.'

that Menard did *not* finish his novel so that the short story in question is first and foremost about there *not* being two *Quixotes* (its very title, *Pierre Menard, Author of the Quixote*, thus designates a non-existent object).

To finish, we can still choose between the image of Menard as a contingent madman in a necessary universe and the image of Menard as a lucid and for this very reason resigned intellect that inhabits a contingent world. No matter which reading we prefer it remains striking that our story treats some of its background themes with a surprising neglect. For instance, the eternal passage of time does certainly not imply that all possibilities will be ultimately realized. Yet Borges lets his Menard remark that 'my undertaking is not difficult, essentially. I should only have to be immortal to carry it out' (this remark is obviously yet another version of Borges' favorite and false quibble that, during the infinite time, every human being will suffer everything). Now there is absolutely no logical reason why this assumption should be true and the few finished chapters cannot serve as a proof here. It is enough to take a look at Menard's other statements about the contingent nature and indifferent value of Cervantes' novel, and at his parallel between his own effort and 'a theological or metaphysical demonstration' in order to realize that a lazy error concerning time and possibility contradicts Menard's most remarkable aim: to make the contingent and the necessary coincide.

BIBLIOGRAPHY

Barnes, Jonathan. 1978. 'La doctrine du rétour éternel' in Brunschwig, Jacques (ed.) *Les Stoïciens et leur logique. Actes du Colloque de Chantilly, 18-22 septembre 1976*. Paris: Vrin: 3-20.

Barth, John. 1977. 'The Literature of Exhaustion' in Bradbury, Malcolm (ed.) *The Novel Today: Contemporary Writers on Modern Fiction*. London: Fontana: 70-83.

Borges, Jorge Luis. 1962. *Ficciones* (transl. Anthony Bonner). New York: Grove Press.

Cicero. 1951. Nature of the Gods – Academics (transl. H. Rackham). Cambridge, Mass.: Harvard University Press.

Compagnon, Antoine. 1979. *La seconde main ou le travail de la citation*. Paris : Seuil.

Krauss, Rosalind E. 1981. 'The Originality of the Avant-Garde: A Postmodern Repetition' in *October* 18: 47-66.

Long, A. A., 1985. 'The Stoics on World-Conflagration and Everlasting Recurrence' in *Southern Journal of Philosophy*, suppl. vol. 23: 13-38.

McGinn, Colin. 1993. *Problems in Philosophy. The Limits of Inquiry*. Oxford: Blackwell.

Plato. 1988. *Republic 10* (transl. Stephen Halliwell). Warminster: Aris and Phillips.

Plato. 1998. *Cratylus* (transl. C. D. C. Reeve). Indianapolis: Hackett.

Unger, Peter. 2002. *Philosophical Relativity*. Oxford: Oxford University Press.

Wilson, Margaret Dauler. 1999. 'Possible Gods' in Id. *Ideas and Mechanism. Essays on Early Modern Philosophy*. Princeton: Princeton University Press: 407-420.

JACQUES MORIZOT

Menard, from Literature to Visual Arts

Abstract: This paper examines a rather neglected topic concerning the Menard case, one that the notationaly-oriented approach leaves aside. Even if Borges describes Menard as a writer, it seems nonetheless illuminating to confront his project with non-literary realizations in the twentieth century. It is not only a welcomed opportunity to refer a fictional character to real persons, it is also a convenient way to find fruitful correlations with some initiatives in the artistic field. If Menard is not a forger, he is likely to be interpreted as a performer or as an appropriationist, both practices that have played a significant role in the recent development of art.
Keywords: textualism; authorial intent; forgery; performance.

Borges's story, 'Pierre Menard, author of the *Quixote*,' has become a favorite and irreplaceable reference for the ontology of art. Innumerable papers have been written to analyze its relevance to the main topics pertaining both to the notion and to the philosophy of art. Among the most significant ones: the status of literature, the aesthetic implications of the similarity between two distinct works, the way of defining a work, the role of authorial intent, and so on. Although Borges is sometimes presenting his short story as the best means he found for testing mental sanity after a serious disease (Borges 1970),[1] it is likely that it does not come down to a theoretical curiosity born in the fertile brain of a writer too fond of possible worlds and narrative puzzles. Under anecdotal setting, it deals with what lies at the core (either fascinating or dreaded) of creative human behavior.

In this paper, I should like to examine a rather neglected topic, namely the insights derived from a comparison between situations similar to Menard's in different kinds of art. As a matter of fact, such a point of view is not only the other side of approaches exclusively based on the textual medium, it also offers

---

[1] Other versions are quoted by Jean-Pierre Bernès in his notes to Borges (see Borges 1993: 1555, 1570, 1696).

a way of connecting fiction with effective practices brought by artists to surprising results.

1. It is a fact that Borges imagines Menard as a poet but one suspects that the real scope of his "secret work" carries indeed far beyond the literary field. Not that some linking with the symbolist or esoteric background are not relevant (the same is equally true for Duchamp, cf. Clair 2000), but the contemporaneity with the Dadaist avant-garde stands a chance of being more decisive than the mere use of words, even if Menard's supreme impudence is to keep his project unknown to everybody.

Of course, nothing prevents anyone from stressing that Menard is writing a text, not anything else, and even from looking for clues in the semiotic peculiarities of linguistic symbol systems. There is nonetheless a paradox along this line of thought, best exemplified by Nelson Goodman's treatment of what he calls the 'Borges Problem' (Goodman – Elgin 1988). This problem stems from the choice of the text as a criterion of identification for literary works. In so far as all languages, natural or artificial, share the characteristic feature of being highly differentiated in their syntactic structure, it is natural to think that the most natural way to capture what is idiosyncratic of every work built out of sentences has to do with the properties of the text in which the work is embedded. But if a work of literature is identified solely on the basis of its text, the unpleasant consequence is that there is no possibility to tell apart two works that would share exactly the same string of signs. So the argument runs: Menard has not written a work of his own, he just adds by an improbable detour a supplementary item to the long list of *Quixote* instances.

This conclusion sounds very odd to the majority of scholars and readers, and I doubt whether it can be supported by anyone who has a strong involvement in literature. So it is not surprising that a practitioner in literary studies like Gérard Genette discards Goodman's thesis altogether, even if he comes to an agreement with him on many subjects and even on the corrective import of his methodology. What's the matter with a textually oriented policy? We can already guess that what is at stake is not only the dismissal of some feature responsible for the unacceptable

conclusion[2], for it would soon eventuate in a mere displacement of the difficulty. The question is worth being seen not so much in the diagnosis of a flaw than as a riddle, that is the discovery that something important is not taken into account and leads to paradoxical results.

It is essential to notice that resorting to the underlying principle: one text – one work does not obtain with the same strength in each direction.[3] Everyone will agree that every alteration in a text engenders a different work, though in practice it may be difficult to determine the due amount of admissible variation. Probably, it would not seem reasonable to assert that a book with a single misprint is not an exemplar at all of the work considered. Goodman himself, who maintains that, due to the transitivity of identity, acceptance of the slightest departure from the original would lead to the effect that all copies might be of the same work (Goodman 1976, esp. pp. 186-187), acknowledges that everyday policies cannot follow such a strict theoretical criterion. The theory of notationality is responsible for this: the result holds as a consequence for a script or a musical score but has no exact counterpart for sketches. The alleged reason is that, in pictures, there is no means to discard any difference as unimportant or, in more technical terms, no possibility to sort pictorial marks into characters. It cannot be forgotten either that this argumentation is dependent on a given cultural background that stresses the importance of genuineness and individual creativity; the impersonality of constructivist sculpture or the recent development of computer art for instance would be claiming in the opposite direction, which leads to relativize the conclusion as a general policy.

As Borges insists that Menard undertakes to write the *Quixote* again: he does not want to write a new version of Cervantes's book but really to recreate it a new personal guise, it is no surprise that, making use of an articulated symbolic system and being finicky about the way to accommodate it, the result is that we are confronted with two identically spelled texts. But at this stage, nothing is said yet concerning its contribution to the work and even to a work.

---

[2] Of course, this does not mean that proposals such as Schwartz (1993) or Davies (1991) are of no value.

[3] For an examination of what is at stake at this point, see Morizot 1999. The present paper may be read as a remainder; it affords additional material, particularly in its second part, and to some extent, is directed towards a reorientation.

In order to give shape to this separation, it may be useful to draw a distinction between sheer repetition and the act of deliberately using repetition. To repeat is simply to produce a second instance similar to another one, without any consideration of method or motivation. Nevertheless, the mere use of repetition induces some perplexities, as the French philosopher Vincent Descombes has casually remarked, in a book whose French version is entitled *The Same and the Other* that offers a broad survey of French philosophy between the 1930s and the 1970s (Descombes 1979). To stress the case of disparity inside the most perfect resemblance (a belated sequel of the never-ending problem of the one and many), he had the idea to add a supplementary page just after the title page. This page was a duplicate of the real title page, except for the fact that the bottom part of it – where ordinarily one finds mention of the name of the publisher – was showing the following inscription:

Foreword to the reader:

This page reproduces the preceding one. *Other*, it is the *same*. But to prevent the reader from counting this second first page for nothing, for instance by attributing it to a binding error, I had to inscribe this foreword that does not appear on the first first page. To be the *same*, it must be *other*.

The particular content of this page is of little matter, the important fact is that resemblance is not so native a relation or rather that awareness of it is in some way revealed by non-coincidence. This result may be of a general scope and echoes the case of the monochrome, that is to say a picture in which the whole of the pictorial content has been neutralized. It seems reasonable to admit that a white monochrome (to stick to a narrow domain) is a very literal picture, since it can be described simply as a rectangle of white paint. But it would be altogether wrong to infer that all monochromes are the same, first because the perceptual evidence is open to minute variations, and secondly because the aesthetic experience makes allowance for the context of creation and interpretation. To see a painting that is white is not to see a white thing in general. In this respect, Danto is undoubtedly right to argue that to see something as a work of art requires something that is beyond the eye's grasp and presupposes some kind of theoretical frame.

Each monochrome tells in fact its personal story. For example, when Malevich paints his *White Square on White Background* (1919), he is aiming at a metaphysical purpose; he doesn't express something of the visible but reveals the need for the vanishing of appearances, which is the requisite to convey reality as energy. By contrast, the so-called *White Paintings* (1951) by Rauschenberg are not an outgrowth but a beginning in his career, just the opposite of Reinhardt's black *Ultimate Paintings*. Although he has explicitly discarded an interpretation in terms of a Dadaist provocation and claimed "a single white as a single God", it must be recalled that these paintings were designed under the influence of John Cage and used in stage settings. Rauschenberg insisted that the pure whiteness gives the opportunity to lay stress on reflections and silhouette shades projected by people passing in front of them. After 1965, Robert Ryman devoted the whole of his activity to painting white pictures. The essential point for him was to exemplify the material process of realising; he denied his paintings lack colours because their main feature was to play with texture, interaction between physical substrate and paint that permits transparencies or fatness of coat, and so on. As Naomi Spector puts it, Ryman's work is "to paint the paint", to rediscover the primitive stages of painting less as an art than as a physical activity.

The lesson involved in these brief insights, unexpected for many, is that the definition of monochrome as a genre does not entail the monotony of its reception as a work or that there is no aesthetic redundancy in this kind of pictures. As a consequence, a work of art of any kind cannot be reduced to a formal scheme. The descriptive success and aesthetic limitations of formalism tended to disconnect a work from its act of creation. But in fact, any work has the properties it has as a result of a web of intentions, either from the artist or the spectator who situates it in a particular context. The textual sameness, which was evidence of identity within a semiotic frame, becomes with the Menard case a clue of originality in his artistic project.

As a consequence, Menard is better viewed as a performer rather than as a writer – I mean he is performing, though not like an actor who has learnt his part, but like an artist who is contriving some happening. If description-of-*Quixote* (hyphenated in Goodmanian fashion) there is, it cannot be explained within the bounds of what is written – it would then amount to

sheer reiteration – but as a game in which the text is but the material.

As could be expected, the best examples are to be discovered among artists, especially visual artists that resort to bookwork, that is artworks in book form.[4] I just want to name here Rodney Graham, who uses the content of some famous works rather as a medium than as a text, strictly speaking. Unlike Menard, he does not hesitate to disrupt the strict sameness of the verbal sequence while obeying the author's logic of writing; the upshot is that he engenders interpolation (Melville, Poe) or endless loops (Lenz, the score of *Parsifal*).[5] Maybe the best instance is *The Veranda*, after Melville's short story, translated and published in Brussels in two booklets tied up by a string of paper. As the subject of Melville's story is the addition of a veranda to his house, Graham, on the pretext that Melville's text is weaved with literary and mythological allusions, adds on behalf of himself an architectural ornament which permits him to narrate the Vitruvian version of the Corinthian capital's origin. On the cover of it, the title has been put into brackets and the author's name replaced by the artist's one. It is up to the reader to draw suitable comparisons between the two texts and infer what is about the relation between text and work.

2. As strange as it may seem, Menard's feat was fully matched by an often-disregarded episode known as the Hank Herron case. In 1973 appeared an anthology of the first writings on conceptual art, Gregory Battcock's *Idea Art*. Besides more or less famous papers, it contained a short article under the title "The Fake as More" – with the signature of a certain Cheryl Bernstein – that presents itself as a critic's review of an exhibition. But the whole thing is indeed a very intricate and subtle matter.

First, Cheryl Bernstein was a pseudonym or rather a fake identity that concealed Carol Duncan.[6] Hank Herron didn't exist either, of course, but was supposed to be very successful in his artistic initiatives. The full content of the mentioned exhibition was coming down to facsimile paintings of Frank Stella belonging to his first decade of creation (mainly stripe paintings). So

---

[4] On bookwork as a genre see Mœglin-Delcroix (1997), on Graham see pp. 312-318.

[5] For other developments on Graham cf. Morizot (forthcoming); also Wall (1988).

[6] The above paper has been reprinted in Duncan 1993, a selection of her personal essays.

we should be tempted to conclude that all this amounts to a trite case of forgery. Sure it is but recognizing it does not exhaust the question.

As Thomas Crow remarks (Crow 1996, ch. 4), the so-called Herron has a genuine prototype in the artist Elaine Sturtevant who, between 1965 and 1973, produced replicates of Stella, Oldenburg and Warhol. As a matter of fact, it happened that the imagined painter superseded the actual one, who resigned painting soon after. It also vividly anticipates what the next decade would go to name "simulation art". This artistic trend comprises lots of tendencies, all of them having however to do something with the idea and practice of appropriation: from Cindy Sherman or Richard Prince who are playing with publicity or movie *clichés,* to the neo-geo movement that critically reinterprets the artists emblematic of modernism[7] and to the virtual annexing of the whole art history by Mike Bidlo. Perhaps the most paradigmatic instance is Sherrie Levine, who was claiming that "we can only imitate a gesture that is always anterior, never original" or that "a painting's meaning lies not in its origin, but in its destination" (Levine 1992: 1067).[8]

In the vicinity of these artists, we seem to be at the same time very far from and very close to Menard's name. For we cannot forget that Borges knows in reality and Menard in the fiction that Cervantes has already written a novel whose title is the *Quixote*. Each of them is supposed to have read it – in fact, read and forgotten as Borges remarks somewhere. Had they not known it, the whole story would be meaningless and Menard in the position of an ape in front of a keyboard. On the other hand, such a strange behavior may have a rationale if we accept to consider that its key does not lie between the lines.

Menard might be seen as one that prefigures the appropriationists of the 1980s, not prompted for sure by the same interests but yet partaking in a common field of action. If suspicion of forgery is in the logic of Duncan's publication, it does not fit either Menard who does not imitate the outcome but is contriving a personal path to achieve it, or the above-mentioned artists who make a critical use of cultural borrowing more than they copy the singularity of a given work. By the way, these prac-

[7] Cf. e.g. P. Halley, Ph. Taafe, R. Bleckmann, etc., who refer to Mondrian, Newman, Riley, etc.

[8] On all these artists cf. Siegel (1988).

tices are contemporary of a society of the spectacle, of media and simulacra; as Foster puts it, "simulation can produce a representational effect without a referential connection to the world" (Foster 1996: 103).

Some convergences with Borges's words are really breathtaking. About Cesar Paladion, a Menard's double who also publishes literary duplicates, Borges writes of one of his books that "nothing is more remote from the homonym book […] that did not reproduce any anterior work"(Borges and Bioy Casares 1970: 19). The decision to start from an existing œuvre is the most important factor, whatever the result may be, either a localized influence or an overall import. The same of course is true of Levine when she takes photographs of already taken photographs, by say Evans or Weston, and not realizes new prints or similar settings. Though the mechanized technique tends to abolish any difference between the two pictures, the resulting one does not qualify as a true facsimile; it remains her choice and retroacts on the supposed original one. As a consequence, Borges – but the pseudo Bernstein as well – claims that the second one is "almost infinitely richer", subtler even if in return more ambiguous.

One could be tempted to discard all this and interpret it as a mere joke but we must also be aware that the pages written by Borges are a unique blend of tale and essay, just like appropriationists' work is a mix of creation and theory. It is in any case not at all evident that either of them is driven by a planned destruction of aesthetics. The basic point is that art is no more the making of works but any act of intervening about them. So the displacement of a work is logically more relevant than the nature of it. If we worry too much about works as closed entities, we may likely have trouble accepting it but it doesn't mean there is not any lesson to derive from it.

In my view, literalism is at the same time a threat and a good omen, for its meaning changes altogether according as one considers the use or the mention of it. For sure, literalism might be found guilty of overlooking the core of what constitutes a work of art if it manipulates the whole stuff blindly; it then comes to be an act of vampirism. Instead, when literalism takes care of what it uses as a means, it happens that it draws attention to the particular content of it, the previously unseen or neglected differences. The trickle-back proves itself to be positive. So we don't have to get obsessed by the thorough identity in the guise

of some works. In this context, Menard's case is just an absurd idealisation of what is currently encountered when somebody sees or reads a cultural object anew.

Besides, consider the analogy of a theatre performance. Here too the text of the play doesn't change and nonetheless the work resulting on stage is always different. It's quite possible the best pictorial equivalent would be critical discourse as taken up by British aestheticians of the 1950s. Among others, Margaret Macdonald explicitly compares the task of the critic with that of the actor or executant. "Critical talk about a work [she writes] is, as it were, a construction of it by someone at a particular time, in a certain social context." If criticism – and even appraisal – is "more like creation than like demonstration and proof", a fruitful continuity is then revealed between artist and spectator and perhaps between the first attempts of the layman and the outcome of the expert glance as well.

Does something really prevent from subscribing to this line of argument? It would be easy to object that there is nothing more than a bad coincidence – or a contingent junction too good to be true – between situations that do not belong to the same world. Maybe one could even suspect that it reiterates the underestimation of context noticed above in textually oriented studies. The difference is nevertheless that the logic of decontextualization is an integral part of these practices but is alien to semiotics altogether. It seems in any case that Borges himself is backing it up, since the last paragraph is explicitly dealing with a form of criticism free from author's constraints.[9] He claims for a calculated use of anachronism, because finally Menard's product is an inexhaustible palimpsest. Everything happens as if the most literal similarity was the necessary condition for the power of recreating to be maximal. Menard is designing a frame of its own, though physically and semiotically it remains invisible or rather undetectable; evidence that the content has been actually modified is conversely a convincing clue to the effectiveness of this frame. Menard's strategy is to some extent at the opposite of that of Tafas (in *Bustos Domecq Chronicles*): the latter paints careful vistas of Buenos Aires then covers them with black polish; doing so, he erases the contents, depriving it from any sig-

---

9 This proposal is at loggerheads not only with a robust theory of literary meaning like that of Hirsch or Stecker but even with the kind of hypothetical intentionalism favored by Levinson (1996).

nificance. Instead of subtracting, Menard is continuously adding; he goes on multiplying corrections and variations in rough, witness the mass of scraps that he filled in … before getting rid of them. The most visible trace is there's no remaining personal mark, but in principle it could be the job of some shrewd critic trained in genetic studies to trace every step back, as Borges hypothesizes.

So my intention is not to uphold what would count as the best available reading. Indeed I believe it is preferable to keep the part of mystery and fascination emitted by these pages away from too firm convictions and foregone conclusions. All I have tried to do in sum is to suggest that it is highly rewarding to put Menard's story in a context or prospect of experimental art researches and not to see it only as a thought experiment. From this perspective, the tale devised by Borges is also a matter for a contractual theory of reference, as Umberto Eco sketches it (Eco 1999, ch. 5). From now on, by tenuous but insistent threads, Menard belongs for better or for worse to art history, at least as much as he belongs to the repertoire of logical puzzles. Each work is a fiction (even here the alleged writer!) though the string of them, together with its halo of comments and perplexities, is part of the fabric of the artworld and of reality.

BIBLIOGRAPHY

Borges, Jorge Luis. 1970. 'An Autobiographical Essay' in *The Aleph and Other Stories, 1933-1969* (tr. N. T. di Giovanni). New York: E.P. Dutton & Co.: 208-223.

Borges, Jorge Luis. 1993. *Œuvres complètes I*. Paris: Gallimard. Bibliothèque de la Pléiade: 467-475.

Borges, Jorge Luis and Bioy Casares, Adolfo. 1970. *Bustos Domecq Chronicles* (tr. F. Rosset). Paris: Denoël.

Clair, Jean. 2000. *Sur Marcel Duchamp et la fin de l'art*. Paris: Gallimard.

Davies, David. 1991. 'Words, Texts and Contexts: Goodman on the Literary Artwork' in *Canadian Journal of Philosophy* 21(5): 331-346.

Descombes, Vincent. 1979. *Le Même et l'autre. Quarante-cinq ans de philosophie française (1933-1978)*. Paris: Éditions de Minuit.

Duncan, Carol. 1993. *The Aesthetics of Power*. Cambridge: Cambridge University Press.

Eco, Umberto. 1999. *Kant and the Platypus* (tr. A. McEwen). New York: Harcourt.

Foster, Hal. 1996. *The Return of the Real*. Cambridge (Mass.): MIT Press.

Goodman, Nelson. 1976. *Languages of Art. An Approach to a Theory of Symbols*. Indianapolis: Hackett.

Goodman, Nelson and Elgin, Catherine Z. 1988. 'Interpretation and Identity: Can the Work Survive the World?' in *Reconceptions in Philosophy and Other Arts and Sciences*. Indianapolis: Hackett: 49-65.

Levine, Sherrie. 1992. 'Statement' in Harisson, Charles and Wood, Paul (eds.) *Art in Theory (1900-1990).* Oxford: Blackwell Publishing: 1066-1067.

Levinson, Jerrold. 1996. 'Intention and Interpretation in Literature' in *The Pleasures of Aesthetics*. Ithaca: Cornell University Press: 175-213.

Mœglin-Delcroix, Anne. 1997. *Esthétique du livre d'artiste (1960-1980)*. Paris: Jean-Michel Place – Bibliothèque Nationale de France.

Morizot, Jacques. 1999. *Sur le problème de Borges. Sémiotique, ontologie, signature*. Paris: Kimé.

Morizot, Jacques. 2004. 'Littéralité et "Menardisation"' in *Interfaces: Texte et image*. Rennes, Presses universitaires de Rennes: 63-75.

Siegel, Jeanne. 1988. *Art Talk: The Early 80s.* New York: Da Capo Press.

Schwartz, Robert. 1993. 'Works, Works Better' in Erkenntnis 38(1): 103-114.

Wall, Jeff. 1988. 'Into the Forest: Two Sketches for Studies of Rodney Graham's Work' in *Rodney Graham, exhibition catalogue*. Vancouver: Vancouver Art Gallery.

KAREL CÍSAŘ
Arthur C. Danto, Author of *The Transfiguration of the Commonplace*

Abstract: Danto's innovative theory of art – formulated in response to artworks that were made by being transfigured from commonplace objects – was to undergo two radical transformations. On the one hand, this theory instigated the rise of the so-called Institutional Theory of art, whose conclusions are rejected by Danto; on the other hand, it produced a theory of the end of art, which he later developed – as a reinterpretation of his original conclusions. Transformative appropriation is the most natural element of Danto's philosophy, as well as of the art of the neo-avant-garde followers of the original avant-garde on which Danto's theories are based.
Keywords: Arthur C. Danto; transfiguration; appropriation; Neo-Avantgarde.

The reasons that lead one to make what at first may appear to be a surprising analogy between Pierre Menard and Arthur C. Danto, can be found in the preface to Danto's book *The Transfiguration of the Commonplace* (Danto 1981: v). If Pierre Menard undertook to write Cervantes' *Don Quixote*, then Arthur C. Danto – at least according to what he claims in the preface to his book – did something similar with the work, *The Transfiguration of the Commonplace*, which was written by a character called Helena in the novel *The Prime of Miss Jean Brodie* by the British novelist Muriel Spark. Unlike Menard, whose work remained a mere fragment (comprising chapters nine and thirty-eight of the first part of *Don Quixote* and a portion of chapter twenty-two), Danto managed to complete his book, *The Transfiguration of the Commonplace*. It corresponds perfectly with the original, even though it consists only in the title – for nothing else remained of Helena's book in *The Prime of Miss Jean Brodie*.

As he has said, Danto actually tried to learn more about the contents of Helena's book by asking the author of her story, Muriel Spark. To his delight, she wrote back to say that the book would have dealt with art in a way that corresponded to Helena's own concerns. According to Danto, her book would

therefore have dealt with the "transfiguration of a commonplace girl into a fictitious character". Conversely, he was to attempt the "transfiguration of a fictitious book into a real one". Danto again refers to the circumstances under which he wrote his most famous philosophical work in the preface to a selection of his critical texts, published in 1994 under the title 'Embodied Meanings' (Danto 1994: 7-8). Here, however, he does not relate them to a transfiguration between the real and the fictitious, but to an internal transfiguration of himself. After writing *Analytic Philosophy of History*, *Analytic Philosophy of Knowledge* and *Analytic Philosophy of Action*, he no longer wanted to write *An Analytic Philosophy of Art*. Instead, he appropriated for his book the less abstract title *The Transfiguration of the Commonplace*, the aim being to reach and convert a wider audience. But it was mainly Danto himself who was converted by this book, for the philosopher was to become an art critic.

1. Danto's book, *The Transfiguration of the Commonplace*, also deals with this specific transfiguration. From the preface we learn that this is nothing other than the transfiguration of a commonplace object into an artwork, regardless of its aesthetic quality. The ready-mades of Marcel Duchamp and, in particular, the work of Andy Warhol are given as examples here. If Marcel Duchamp was the first to turn a bottle-rack, bicycle wheel or urinal into artworks, it was Andy Warhol who, in Danto's view, turned Duchamp's gesture into the actual subject of his artistic activity.

Danto makes several references to his first encounter with Warhol's work in a 1964 exhibition at the Stable Gallery in New York as being a turning point in his philosophical career (Danto 1981: vi-vii, 1994: 6-7 and 1992: 5-6). When he saw Warhol's Brillo Box, at first glance piles of ordinary soap powder boxes, he realized that, for the first time, a genuinely philosophical question had arisen from the art world, namely how can we view these boxes as artworks, which is what his Brillo boxes undoubtedly were considered to be.

Danto's answer, as formulated initially in the articles 'The Art World' (Danto 1964) and 'Artworks and Real Things' (Danto 1973) and subsequently in his book *The Transfiguration of the Commonplace*, is well known. The Brillo boxes cannot be revealed as artworks through mere observation. To be seen as such, it is necessary to have a certain knowledge of art theory and history and to be engaged in the art world. Only after

adopting such an aesthetic approach can one see – albeit on the basis of perceiving what is indistinguishable – not piles of containers for soap pads but artworks. We cannot comprehend a work of art only by looking at it; it has to be interpreted in a kind of strategic way that is oriented in the art world. We have to set out a theory concerning the meaning of the work. It is such an interpretation that transfigures everyday utilitarian objects, such as cartons for soap pads, into artworks.

Danto's innovative theory of art – formulated in response to artworks that were made by being transfigured from commonplace objects – was, itself, to undergo two radical transformations, which are described at the end of the preface to *The Transfiguration of the Commonplace*. On the one hand, this theory instigated the rise of the so-called Institutional Theory of art, whose conclusions (as set out by Richard Sclafani and George Dickie) are rejected by Danto; on the other hand, it produced a theory of the end of art, which he later developed – as a reinterpretation of his original conclusions. Danto criticizes the Institutional Theory, which made famous his concept of the art world, for changing the position used when formulating the underlying issue. Danto was originally engaged in what was basically a Kantian question about the conditions governing the possibilities of artworks; in contrast, institutional theorists posed a different question, i.e. how can a mere thing become an artwork, which they answered, at sharp odds with Danto's theory, in a non-cognitive way, i.e. that a mere thing becomes an artwork once it is appreciated as such by the art world.

In the preface to his book, Danto outlines the distinct reinterpretation of his original theory that he fully developed three years after the publication of *The Transfiguration of the Commonplace* in his article 'The End of Art'. Here, he only notes that, unlike previous definitions of art, his specification is capable of comprising also a revolutionary transfiguration of art of the kind that was brought about by Warhol. Moreover, in Danto's view, this represents a culmination of art history, for it poses a question about the essence of art which, in a Hegelian way, is transfigured into its own philosophy

Apart from differentiating Danto's position from that of the institutional theorists, *The Transfiguration of the Commonplace* also presents objections to the position of the Institutional Theory itself. According to Danto, the main problem with this theory is that it continues to adhere to aesthetic qualities as a con-

stitutive element for the creation of an artwork, a position from which Danto's own theory radically differs. In this connection, Danto cites Dickie's claim that Duchamp's 'Fountain' (a commonplace urinal) displays similar qualities, for example, to sculptures by Brancusi or Moore, i.e. an oval shape and shiny surface that reflects the surrounding space (Danto 1973: 93). Such qualities, however, are displayed by any other urinal – regardless of whether it is considered to be an artwork. For this reason, it is not possible for the art world to appreciate Duchamp's 'Fountain' as an art work for these qualities.

Furthermore, according to Danto, the Institutional Theory cannot properly specify either who is a member of the art world or how one can become a member. For in 1964, in relation to the Brillo Box, according to the Institutional Theory, the philosopher Arthur C. Danto would be included as a member of the art world, as opposed to indisputable art experts such as the director of the Stable Gallery, Eleanor Ward, or the director of the Canadian National Gallery. Charles Comfort, who, at the time, refused to appreciate the Brillo Box as an artwork (Danto 1992: 36-37).

It was precisely this experience that led Danto to adopt the familiar position of Heinrich Wölfflin, according to whom the definable basis of art does not change but is interpreted in considerably different ways in various historical periods. What is now seen without difficulty as art could not have been seen as such earlier on. Moreover, according to Danto, Warhol's Brillo Box exhibition at the New York Stable Gallery in 1964 can, from a certain perspective, be seen as representing the end of art, for in it – by questioning its very essence – transfigures art into its own philosophy. Danto's book contains three references to the theory about the end of art, each time in the context of Hegelian influence. Hegel's philosophy of history is basically correct: the development of art did not come to an end, but to a logical climax: art became self-conscious in a philosophical theory about itself (Danto 1981: vii-vii, 56-57, 111).

The above theory was fully developed in the 1984 essay on 'The End of Art' (Danto 1984: 5-35). Even here, Danto's perspective is based on the historical conditionality of art activity and a Hegelian conviction as to the progressive nature of its development. This may be illustrated by the gradual sophistication of the visual reproduction of reality by means of visual art, which was dropped only with the full development of the narrative means of cinematography, which gradually took over this

role from visual art at the beginning of the twentieth century. The alternative is a theory that subordinates the mimetic function of art to one that is expressive. Such a theory, however, cannot satisfactorily interpret the development of art history as a continuum, for it holds individual artworks to be a mere expression of the artist's disparate and isolated emotions. Against these two theories, therefore, Danto sets a different development model based on the Bildungsroman, that is a novelistic description of the narrator's character development. By applying schemata from Hegel's Phenomenology of Spirit, Danto concludes that essentially self-reflexive philosophy is the culmination of the spirit and art. According to Danto, this is borne out by the development of art in the last few years, in which theory has been increasingly shown to replace the actual artwork. Art has ended, for it has transformed into its own philosophy. From now on, all artworks may be considered to be post-historical works, for Warhol's 'Brillo Box' deprived them of their historical role in 1964.

2. The possibility of there being two observationally identical artworks – like the 'Brillo Box' where the artwork is perceptually indistinguishable from a commonplace object – was, according to Danto, first described in Borges' story 'Pierre Menard, Author of The *Quixote*' (Danto 1981: 33-39). This story is about two completely identical texts of *Don Quixote*, one of which was written by Cervantes, the other by the early twentieth century symbolist poet Pierre Menard. In Danto's view, this raises an ontological question about the identity of an artwork. The answer, however, cannot be found in the perceptual qualities of the two works, for they are observationally identical. Danto finds support for his assumption that they are different with Borges' narrator, according to whom the *Quixote* of Menard is infinitely richer that Cervantes' ponderous *Quixote*. The explanation for this difference, however, cannot be found in the internal qualities of the text; one has to look at the external qualities which place the text in a specific context. For the texts were written at different times and by different authors whose nationalities and aims also differed. According to Danto, their place in the history of literature is also an essential part of these works.

Although Danto's interpretation emphasizes the identity of both works, his intention from the outset is, by contrast, to underline their difference. Besides, he cannot find support for his

claim as to textual identity in Borges' story, for Menard's undertaking is described in several places as being impossible to carry out. What we have in front of us are not two identical texts, but the complete original work of Cervantes and several fragments of this work written by Menard. So it is hardly about an identical work. As far as Danto's interpretation is concerned, the intellectual experiment mentioned by Borges' narrator should be seen as a productive rather than unsuccessful attempt by Pierre Menard. For the narrator admits to being able to read Cervantes' *Quixote* as if it had been written by Pierre Menard. Adding such a footnote necessarily transforms the meaning of the text, turning it into a different work. As mentioned above, it is Borges' narrator who declares Menard's *Quixote* to be richer than Cervantes' ponderous *Quixote*. In addition, it is the narrator who says that Menard "(perhaps without wishing to) has enriched" the rudimentary art of reading by means of a technique of deliberate anachronism and erroneous attributions, which makes it possible to interpret any text as having been written with different authorial intentions and in different historical contexts.

Perhaps what brings together the positions of Danto and Borges' narrator more than anything else, is that they both transfer the initiative for finding the meaning of a work from the author to the reader or observer. As is the case with Warhol's 'Brillo Box' – where Danto does not enquire as to Warhol's intention, but identifies it from an observational position – in Borges' story, too, an evaluation of the meaning of Menard's work is left to the informed reader. Danto proceeds in a similar way in other cases covered in *The Transfiguration of the Commonplace*. Perhaps the most famous example is his intellectual experiment involving a group of identical red square canvases, each having a completely different meaning and, in certain cases, a different ontological status. Depending on the external qualities of the work – e.g. its title, the author's personality or the work's place in art history – they may represent a historical painting, a psychological portrait, a landscape, a geometric abstraction, religious art, a still-life, or a red object that is not connected to art in any way. Everything depends on the ability of the observer to make use of these external leads and, on their basis, to formulate a theory about the meaning of the work.

Although this possibility is treated only peripherally in *The Transfiguration of the Commonplace*, Danto's theory about the difference of perceptually indistinguishable artworks need not be

verified solely by considering intellectual experiments. They can be proved by viewing the actual artworks, in the way the author himself undertakes to do in his later texts. The relevant examples may initially be divided into two groups: monochrome paintings (artworks that reduce the means of expression to a single quality of colour) and ready-mades (artworks made by being transformed from commonplace objects).

In the book *The Transfiguration of the Commonplace* Danto focuses on the work of Ad Reinhardt and tries to show, above all, the uniqueness in style of Reinhardt's monochrome paintings (Danto 1981: 204); in his later monograph *Beyond the Brillo Box* he focuses on the theory about the end of art and considers the differences between the otherwise indistinguishable monochromes of Marcie Hafif (1981), Ad Reinhardt (1962) and the first ever monochrome paintings undertaken by Kazimir Malevich in 1915 (Danto 1992: 46-47). Paradoxically, an interpretation of later neo-avantgarde monochromes, which for Danto are only a post-historical repetition of an original avant-garde gesture setting the boundaries of art, is essential to prove his early theory about the role of a work's external qualities when finding its meaning. For it is on these monochromes that Danto can convincingly show that a work of art necessarily relates to its precursors, and is even a kind of criticism. The pioneering works of Malevich and Rodchenko, however, are more difficult to interpret; in fact, they defy interpretation – at least from the perspective of the artists. In their radicalism they reject the traditional metaphoric structure of an artwork, upon which Danto explicitly leans in *The Transfiguration of the Commonplace*. Rodchenko's famous triptych 'Pure Colours: Red, Yellow, Blue', dating from 1921, is a good example of this. According to Rodchenko himself, the aim of the work is, on the one hand, to rid colour of any denotative function and, on the other, to mechanize the process of painting. With this, however, disappear two fundamental leads for identifying the meaning of a work, as formulated by Danto in *The Transfiguration of the Commonplace*. Retrospectively, it is possible to identify a place for Rodchenko's triptych in the art world, but such a place is different to the one originally claimed by the work.[1]

A similar case applies to Danto's interpretation of ready-

---

[1] For the relationship between avant-garde and neo-avantgarde monochrome paintings, see Buchloh 1986: 41-52.

mades. In *The Transfiguration of the Commonplace*, Danto focuses on Andy Warhol, but in his later texts there are more references to the specific form of ready-mades, which is appropriation of another artwork (and not a commonplace object) (Danto 1994: 58). For Danto's early theory about the difference of perceptually indistinguishable artworks, similarly to monochrome paintings, it is less difficult to interpret Warhol's later ready-mades or the appropriations exhibited, for example, by Sherrie Levine.[2] Levine became known mainly due to her photographic reproductions of works by Edward Weston ('After Edward Weston 1', 1980) and Alexander Rodchenko ('After Alexander Rodchenko 3', 1987); in addition, however she has also undertaken an appropriation of Duchamp's Fountain – a copy consisting of several identical bronze casts. Moreover, Levine is one of the few artists to make explicit references in her work to Borges' story 'Pierre Menard, Author of the *Quixote*'.[3] The meaning of Sherrie Levine's works is actually focused on their external qualities, in particular the artist's female name and the title, in which a key role is always played by the ambiguity of the word 'after' (following in time and in imitation of). Only with these leads is it possible to differentiate her works from their perceptually indistinguishable models and to successfully interpret them as subversions – often openly feminist – criticism.[4]

Duchamp's Fountain is harder to interpret from the perspective of Danto's theory. Unlike Sherrie Levine's reproduced bronze copies, which subversively transform the original ready-made into a traditional sculpture in a traditional style, the status of Duchamp's work is questionable. It is well known that Duchamp's 'Fountain' is actually lost – there are now only copies of the original and one photograph taken by Alfred Stieglitz. Duchamp sent his Fountain – a commonplace urinal, furnished only with a signature ('R. Mutt') and a date ('1917') – to the first annual exhibition of the Society of Independent Artists, of which he was a member. Although it fulfilled the only condition specified in the third section of article three of the society's regulations, according to which the exhibition may involve the

---

[2] For the relationship between avant-garde and neo-avantgarde ready-mades, see Buskirk and Nixon 1996.

[3] Levine 1987: 92-93. See also Wallis 1984, which contains a reprint of a translation of Borges' story.

[4] Cf. Crimp 1993: 126-136; Krauss 1986: 151-170. For the legal aspects of appropriation, see Buskirk 1992: 82-109.

participation of anyone who sends the required fee of six dollars, Duchamp's Fountain was not accepted for the exhibition. Only much later was it appreciated as art, as were Duchamp's previous ready-mades, which successfully resisted being classified as artworks not only due to their radical limitation of their denotative functions but also due to their anonymity and means of display (Duve 1996: 89-143).

It is not Warhol's 'Brillo Box' or Ad Reinhardt's black canvases, but Duchamp's ready-mades – "still almost not art" – and Rodchenko pure colours – "almost no longer art", that represent the original demarcations of the boundaries of art. As such, however, they defy both of Danto's theories. On the one hand, they shake off the external leads by which they can be identified as art, and, on the other, they foreshadow Danto's end of art. Danto's theories are based on an interpretation of the neo-avantgarde followers of the original avant-garde. Transformative appropriation is the most natural element of Danto's philosophy, as well as of their art. This is also why the author of *The Transfiguration of the Commonplace* is Arthur C. Danto.

BIBLIOGRAPHY

Buchloh, Benjamin H. 1986. 'The Primary Colors for the Second Time: A Paradigm Repetition of the Neo-Avant-Garde' in *October* 37: 41-52.

Buskirk, M. and Nixon, M. (eds.). 1996. *The Duchamp Effect*. Cambridge, MA, London: MIT.

Buskirk, Martha. 1992. 'Commodification as Censor: Copyrights and Fair Use' in *October* 60: 82-109.

Crimp, Douglas. 1993. 'Appropriating Appropriation' in: Crimp, Douglas (ed.) *On the Museum's Ruins*. Cambridge (Mass.), London: The MIT Press: 126-136

Danto, Arthur C. 1964. 'The Art World' in *Journal of Philosophy* 61(19): 571-584.

Danto, Arthur C. 1973. 'Artworks and Real Things' in *Theoria* 39: 1-17.

Danto, Arthur C. 1984. 'The End of Art' in: Lang, Beryl (ed.) *The Death of Art*. New York: Haven Publishing: 5-35.

Danto, Arthur C. 1981. *The Transfiguration of the Commonplace*. Cambridge (Mass.): Harvard University Press.

Danto, Arthur C. 1992. *Beyond the Brillo Box. The Visual Arts in Post-Historical Perspective*. New York: Farrar Strauss Giroux.

Danto, Arthur C. 1994. *Embodied Meanings. Critical Essays and Aesthetic Meditations*. New York: Farrar Strauss Giroux.

Duve, Thierry de. 1996. *Kant after Duchamp*. Cambridge, MA: MIT.

Krauss, Rosalind E. 1986. 'The Originality of the Avant-Garde' in: Krauss, Rosalind E. (ed.) *The Originality of the Avant-Garde and Other Modernist Myths*. Cambridge, MA: MIT: 151-170.

Levine, Sherrie. 1987. 'Five Comments' in Wallis, Brian (ed.) *Blasted Allegories: An Anthology of Writings by Contemporary Artists*. Cambridge, MA: MIT: 92-93.

Wallis, Brian (ed.). 1984. *Art After Modernism: Rethinking Representation*. New York: The New Museum of Contemporary Art.

ROBERTO CASATI

Play it:
the Replay Theory of Music Experience

Abstract: I present the replay theory of musical artworks, according to which, when we listen to any performance of a piece of music, we take the piece we listen to as being the very same token, accessed in two different moments in time. I discuss some objections to the theory.
Keywords: identity of musical artworks; replay theory.

At Rick's, Ilsa begged Sam to play it once – as she wanted to hear it again.[1] She did not ask him to play another one of the same type.[2] She did not beg him to play *a* 'As Time Goes By'. We understand what she meant; and so did Sam, who played *it*; and so did Rick, who said he thought he had told him never to play it. Later on, Rick had second thoughts and summoned Sam: "You played it for her, you can play it for me."

Was Sam playing one song, or many? It is usually assumed that there is a general problem with the identity conditions of works of art, most famously epitomized by duplication fantasies such as the one described in Borges' *Menard* fable. The problem consists in specifying satisfactory conditions that allow for the dismissal of dubious cases. For instance, a 'pure' conception of works of art as abstract entities (types) (Kivy 1983), having no essential link to the history of their production, and hence to the life of certain tokens, would endorse counterintuitive cases,

---

[1] In Michael Curtiz's *Casablanca* (1942), Ingrid Bergman never spoke the apochriphous line "Play it again Sam", which gave the title to a 1972 Woody Allen movie. She said : "Play it Sam. Play 'As Time Goes By'". But *again* it was. Sam had played it earlier for Ilsa and Rick in Paris at *La belle aurore*.

[2] Cfr. Jackendoff, R., 1983, *Semantics and Cognition*, cap. III n. 14, p. 246: "We say of the Eroica 'I heard it on the radio last night', not 'I heard one of them (a member of the set of performances) last night'. This suggests that we intuitively think of the Eroica as a single #entity# (a complex #sound# or #group of sounds#) that may be heard on various occasions – just as, for instance, the morning star is a #thing# that may be seen on various occasions".

such as the Menard duplications. On the other hand, a 'pure' conception of works of art as concrete individuals (tokens) (Goodman 1968, Predelli 2001, for a review, Kania 2008) would run counter to our intuitions concerning the repeatability of at least some of those works. These conflicting intuitions may just reflect the fact that we possess cognitive capabilities for identifying and re-identifying individual tokens along with capabilities for recognizing types, and we are at a loss when both kinds of capabilities are mobilized for entities such as works of art. Whatever they are, works of art appear to be somewhat remote from the kind of individuals the tracking and recognizing capabilities evolved to cope with in the first place. Hence the puzzles.[3]

But maybe there is no *general* problem. Different kinds of artworks could be treated in completely different ways and we would be happy to settle agreements on a case-by-case basis. This is what I am going to argue for in what follows. The strategy is very simple. I shall show that we have the resources for such case-by-case settlements, at least in one domain, the domain of musical artworks.

The idea is to take Ilsa at face value. Of course, when we listen (at different, non-overlapping times) to two exactly matching performances of 'As Time Goes By', we listen to *two* items. There are two individuals here, two events or processes[4], standardly considered as exemplifying the same process type. This is clean metaphysics. What about the content of our perceptual experience? The claim I'd like to explore is that, as a default, it is as if we listened to *one and the same* temporal token. The content is as of the same song, literally the same song that we listen to now and we listened to earlier – as if we could go back to the past and run a temporal segment of reality in playback. Here is the partial visual analogy on offer: Re-listening to a song is like revisiting an object, possibly from a different point of view. We may wonder whether the analogy clarifies the point. Insofar as it does, then analogies such as this indicate that we have the conceptual resources for making sense of the idea that

---

[3] The puzzles could then just be a projection on ontology ('dubious status' of works of art) of some misuse of cognitive systems, or poor understanding of them.

[4] The distinction between events and processes is not very important here. We may consider events as point-like, processes as extended temporal entities. What matters, is their difference from material objects such as stones or chairs.

it is numerically the same spatio-temporal individual (and not just the same type) that we listen to when we listen to a song at different times. No matter whether this idea is true to the clean metaphysics of artworks.

The claim is that within our conception of musical entities, insofar as it is shaped by some features of our musical experience, the distinction between type and token is not so sharply salient and, possibly, is not required, given that alternative possibilities are available. Which kind of claim is this? It is not a metaphysical claim; for instance, I am not arguing that the identity of musical artworks is somewhat dependent upon the mind. On the other hand, the claim can be construed as metaphysical, more precisely as a descriptive metaphysical claim about the world as experienced in listening.

1. ILSA AND THE SONG

The key notion here is the perception of processes, and particularly the perception of the same process (the sense of 'same' is of course under dispute here.)[5] Ilsa listened twice to 'As Time Goes By' (ATGB), as played by Sam; the first time she did, the song started at midnight in Paris, on June 11, 1940; the second time it went on at late evening in Casablanca, on December 2, 1941. We hesitate between claiming that she listened to[6] one or to two songs. The choice will standardly depend on considering the song as a token ('Two') or as a type ('One') respectively. Contextual factors decide which interpretation we ought to give to the question. But the type-answer is relatively consensual. It is based on a simple argument. The argument starts from the intuition that (1) it is a bit philosophical to claim that as Ilsa first listened to ATGB, and later listened to ATGB, she listened to two songs. This sounds as strange as claiming that she met two people, as she first met Rick in Paris, and a few months later she met Rick in Casablanca. (2) But now, the only candidate for an entity that remains the same on both episodes of listening is the *type* of song. Hence, (3) asserting that Ilsa listened to the same song amounts to asserting that she listened to the same song type.

Premise (2) is my target.

---

5 Musical experience is best described by experimental psychology, but here I shall rely on some phenomenological features of listening.

6 One can claim that, *mutatis mutandis*, what follows applies to song production as well. But I would not endorse this hypothesis without further discussion.

2. OBJECTS AND PROCESSES

Let us make one more assumption to clear the way from a few possible distracting objections. Assume that sounds are not phenomenal objects, but events in sounding objects or in a medium (for a defence, see Casati and Dokic 1994). Under the assumption, where is the difference between seeing Rick twice and listening twice to ATGB?

Rick continued to exist when Ilsa was not perceiving him. And when she perceived him, in Paris and in Casablanca, he existed in its entirety in front of her.[7] When dealing with the re-identification of material objects, Strawson (1959) required that we need a space or space-like frame that could host non-perceived individuals between two episodes of identification, between two encounters with the same individual. This possibility is of course not available for processes, unless it is two phases of an enduring process that we encounter at two different times.[8] ATGB was nowhere to be heard when Sam was not playing it, between the two episodes in which she listened to it. And when she listened to the song in Casablanca, she was not listening to the second phase of something the first phase of which she had heard in Paris. But how can we claim it was numerically the same process she heard both in Paris and in Casablanca?

Even if we had sufficiently clear-cut and powerful identity criteria for processes, the methodological move not to work within the scope of clean metaphysics would prevent us from using them. Hence we are left with the somewhat simpler task of laying down the reasons that may lead to the belief (however false) that a certain process is the same process as the one we previously listened to.

Hence, proceed backwards in the conceptual order, assume Ilsa *did* listen to numerically the *same* thing when she listened to ATGB first in Paris, then in Casablanca, and ask what would be required for making sense of this situation.

In spite of many a deep difference between processes and material objects, there are at least some superficial analogies

---

[7] Perdurantism, I surmise, is the default naive metaphysical option for material objects.

[8] The best analogy, in the processual realm, for a big object's being visible from different location is a process that begins in one location and ends in the other. In both cases, it is *parts* of the process, and of the process, which are seen/heard respectively. It is not the *whole* of them.

between the perception of a process and the perception of a material object. If you look at Rick, you perceive it thanks to a certain 'shape' of the distribution of reflected light – a certain shape of the optic array. Rick dances and rotates on itself, so that after a rotation it shows to his face, as he did before the rotation, at the end of the sequence the optic array is structured in the same way as it was at the beginning. Moreover, two perceptually indistinguishable people are going to produce the same structure of the optic array. Someone could swap them and we may not notice the swap. We do not have internal criteria – based on the inspection of perceptual appearances, or of the optic array – that could decide the question whether we saw one or two people. If we hadn't the notion of material objects as independently existing in a hidden space when we do not perceive them, the question itself would be meaningless – as phenomenalists would claim. But we do have such a notion: the default of Rick-like objects perception is non-phenomenalist. Granted granted that we are unable, on purely phenomenal grounds, to distinguish the object we are looking at from the process we looked at earlier, we perceptually assume by default that it is one and the same object we are looking at. Seeing Rick-looking entity in Paris, and seeing a Rick-looking entity in Casablanca, is by default conceived of as seeing one and the same token, accessed in different encounters.

3. THE REPLAY THEORY

My suggestion is that the default for auditory perception of processes is quite close to (although not coincident with) a non-phenomenalist position about the visual perception objects. Granted that we do not apply strict criteria for the identity of processes, and granted that we are unable, on purely phenomenal grounds, to distinguish the process we are listening to from the process we listened to earlier, we perceptually assume by default that it is one and the same process we are listening to. Listening to ATGB in Paris and listening to ATGB later on in Casablanca is conceived of as listening to one and the same token, *accessed* in different moments. It is as if we replayed an event in the past, as if we revisited that event. Call this the Replay Theory.

4. TESTING THE REPLAY THEORY

Clarifying some details will help get a better picture of the Replay Theory. This can be done by addressing some of the typical conundrums about the identity of musical artworks. As a matter of fact, if the Replay Theory is right, then the conundrums turn to be artefacts of the obsession with the type/token distinction.

(a) How long does the thing I listen to that I call 'As Time Goes By' *last*? Does it last an average of all its occurrences, or any other measure we can come up with? This is an interesting metaphysical question that affects many theories of the identity of artworks. Do song types themselves have lengths, or is the length of a song somewhat related to that of its performances? If the former is the case, can we make sense of a slow performance as a performance of the same song which we heard in a fast performance? If the latter, how are we to choose between lengths?

In our context, the answer is simply that the length is *revealed* in the performance, without coinciding with the length of the performance, or with any pre-assigned length. Performances are like viewpoints on the single process we listen to: the length can *appear* different in different performances, but it is one and the same length. Of course, we may never be able to claim, of any given length, that it is *the* length of the song. But this does not imply that such a length does not exist.

(b) When does ATGB itself begin? At midnight, at late evening, anywhere else in time? Again, a type/token related problem, which here has a simple solution. Wherever ATGB starts, we happen to listen its starting point *once* at midnight, and *another time* at late evening. For the visual analogy: If we look at Rick in Paris and in Casablanca, where does it begin in space? It depends upon where Rick is when we look at it.

(c) Stretching the theory a bit. Imagine that ATGB is played in Paris, and simultaneously in Casablanca. One or two songs? The Replay Theory answers: One, which is heard in Paris and simultaneously in Casablanca. Here the asymmetries with the case of the visual perception of Rick are all the more apparent. Rick[9] cannot be spotted simultaneously in Paris and in Casablanca. However, an ingenious system of mirrors could be devised so that we could have the impression of seeing the same person at two different locations. One of the two appearances

---

[9] Unless he is a is very big giant; but see the previous footnote.

would 'command' the appearance at both locations, and this again would be a difference from the case of the song.

(d) Imagine that ATGB is played on two different players on two different pianos, with a five-second lag between the performances. One or two? Exclude the idea that Ilsa is listening to one single messy melody. One, the Replay Theory is bound to claim. And it provides a visual analogy, the case of seeing double. Seeing double is seeing *one* thing, only seeing it twice; it is not seeing a single, messy object.

(e) Suppose Sam played ATGB slower, in a different tone, backwards, etc. Is it numerically the same entity Ilsa listens to in each case? Suppose the music is played once by Dooley Wilson[10] and once by Oscar Peterson. Is it numerically the same song? Once more, the Replay Theory says yes. And once more, visual perception is a powerful source of metaphors. You can see the same object in good and bad viewing conditions, through colored goggles, through a glass darkly, reversed in a mirror, and painted by Monet or Licthenstein (as it happened with Rouen's Cathedral). The conductor's baton is like a medium that allows you to selectively discern certain parts and properties of a song.

5. SONGS AND IMAGES OF SONGS

Suppose Ilsa liked ATGB so much, she recorded Sam's performance. Now she listens to the performance on her tape-recorder. Is she listening to the *same* ATGB she listened to at Rick's? On this question the Replay Theory can be neutral, and appeal to one more visual analogy. Ilsa liked Rick, and took a picture of him. When she looks at the picture, is she looking at Rick? An answer to the question about the picture is an answer to the question about the recording, as the playing of a recorded performance is an image of the song (Casati and Dokic 1994 defended this idea). Listening to a recording is like looking at an image. Hence, if we accept that we can see people by looking at their pictures, we should accept that we can listen to songs by listening to their recorded performances.

[10] Wilson, *Casablanca*' Sam, was not a piano payer but a drummer and the music had to be dubbed in by an unknown studio pianist.

6. REPLAYS ARE NOT MEMORIES

It is important to get a proper understanding of replays. When Ilsa listened to ATGB the second time, it is for her as if a past temporal process (or relevant portions thereof), the playing of the song by Sam, was replayed in its entirety. It is the worldly *process itself* that is thought of as replayed, not the *experience* Ilsa had when she first listened to ATGB. Said otherwise: listening to ATGB for the second time is not, for Ilsa, tantamount to *remembering* ATGB – although of course some memories of the first performance can accompany her second encounter with ATGB. Moreover, it is not the listening that is repeated: the second listening episode is a different listening episode.

Ilsa has an experience closer to a *simulation* of a past event – except that the simulation is performed by the world itself – in the case at stake, by Sam. This could be, for Ilsa, a sort of passive simulation, not subject to the will except possibly for the initiating event (her asking Sam to play it).

CONCLUSIONS

Events and processes are not material objects: the asymmetry between them is here to stay – even at the commonsensical level. But the asymmetry, or relevant aspects of it, is overruled by analogies in the case of the perception of songs, of clearly categorizable processes. According to the Replay Theory of musical perception, we conceive of songs as of individuals - spatiotemporally articulated entities that we can visit and revisit in time, very much as we can revisit Rick. We access both people and songs through the unfolding of their appearances. In the case of songs, there is no underlying substance that is required to perdure between encounters. But this difference only matters for clean metaphysics: given general principles that relate re-identifiability to unperceived existence and unperceived existence to substantial existence in non-accessed regions of space, and at the same time deny events any substantiality, we may very well reject the idea that events can be so re-encountered after a while. However, I suggested that we have plenty of conceptual resources, based analogically on the case of the perception of material objects, to construe the perception of a song as the perception of a single, concrete individual, and hence to overcome the rather discomforting mechanical application of the type/token dichotomy and the artificial conundrums it generates for our conception of songs. The question is then open, on what

grounds the Replay Theory is to be preferred to the type/token theory. If we are to remain within the limit of descriptive metaphysics, without making any normative claims about the nature of musical artworks, the Replay Theory has the advantage of being free from some of the conundrums that affect the type/token conception of artworks.

The Replay theory makes an empirical claim about our default conception of songs. In so doing, it suggests that we have the cognitive resources for dealing with songs in a way that respects their token individuality across repetitions. As an empirical theory, it has to be rendered compatible with what is known about the architecture of the mind. The type/token theory was in its own way compatible with some features of the architecture of the mind, in particular with the existence of cognitive systems dedicated to individual re-identification and to type recognition. The Replay theory could be thought of as using resources from these very systems and, on top of them, from (parts of) systems devoted to simulation.

BIBLIOGRAPHY

Casati, Roberto and Jérôme Dokic. 1994. *La philosophie du son*. Nîmes: Chambon.
Goodman, Nelson. 1968. *Languages of Art.* Indianapolis: Hackett.
Jackendoff, Ralf. 1983. *Semantics and Cognition.* Cambridge, Mass.: MIT Press.
Kania, Andrew. 'The Philosophy of Music' in Edward N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy (Fall 2008 Edition)*, <http://plato.stanford.edu/archives/-fall2008/entries/music/>.
Kivy, Peter. 1993. *The Fine Art of Repetition: Essays in the Philosophy of Music*, Cambridge: Cambridge University Press.
Strawson, P.F. 1959. *Individuals*. London: Methuen.